

UNIVERSIDAD NACIONAL DEL ALTIPLANO

**FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA,
ELECTRÓNICA Y SISTEMAS**

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS



**“MODELO PREDICTIVO BASADO EN MINERÍA DE DATOS PARA LA MEJORA EN
LA TOMA DE DECISIONES DEL DEPARTAMENTO DE CIRUGÍA DEL HOSPITAL
REGIONAL MANUEL NÚÑEZ BUTRÓN”**

TESIS

PRESENTADO POR:

MARIZELL COPA MAMANI

SONIA VERÓNICA CHAVEZ VIZA

PARA OPTAR EL TÍTULO PROFESIONAL DE:

INGENIERO DE SISTEMAS

Puno - Perú

2015

UNIVERSIDAD NACIONAL DEL ALTIPLANO
FACULTAD DE INGENIERÍA MECÁNICA ELÉCTRICA, ELECTRÓNICA Y
SISTEMAS

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS

“MODELO PREDICTIVO BASADO EN MINERÍA DE DATOS PARA LA MEJORA
EN LA TOMA DE DECISIONES DEL DEPARTAMENTO DE CIRUGÍA DEL
HOSPITAL REGIONAL MANUEL NÚÑEZ BUTRÓN”

TESIS PRESENTADA POR:

MARIZEL CCOPA MAMANI.

SONIA VERONICA CHAVEZ VIZA.

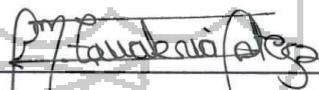
PARA OPTAR EL TÍTULO PROFESIONAL DE INGENIERO DE SISTEMAS

APROBADA POR EL JURADO REVISOR CONFORMADO POR:

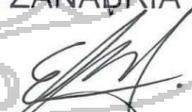
PRESIDENTE


MG. ROBERT ANTONIO ROMERO FLORES

PRIMER MIEMBRO


M.SC. MILDER ZANABRIA ORTEGA

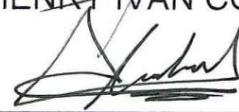
SEGUNDO MIEMBRO


ING. EDWIN FREDY MAMANI CALDERON

DIRECTOR DE TESIS


DR. HENRY IVAN CONDORI ALEJO

ASESOR DE TESIS


ING. ALAIN MELQUIADES ALEJO HUARACHI

ÁREA: Informática

TEMA: Sistemas de información tradicionales y expertos

AGRADECIMIENTOS

A Dios por permitirnos vivir y darnos el entendimiento necesario para llegar hasta el final del trabajo de investigación.

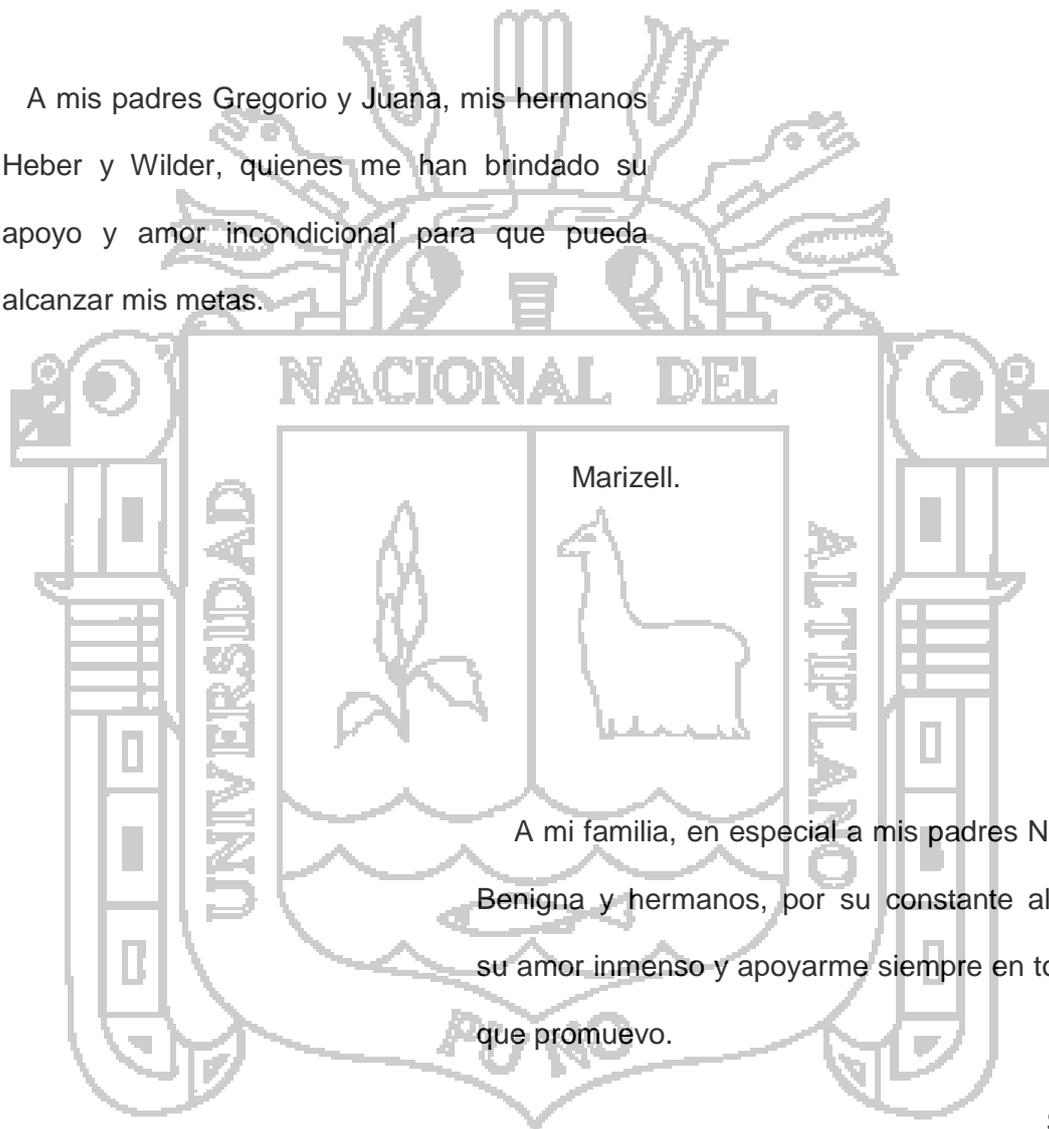
A nuestro Director Dr. Henry Condori Alejo, Asesor Ing. Alain Alejo Huarachi y a todos nuestros jurados por todos sus consejos y su apoyo incondicional en la elaboración de la investigación.

Al personal de Salud del Hospital Regional Manuel Núñez Butrón en especial al Jefe de la Unidad de Estadística e Informática Ing. Domingo Rivera Mendoza, por su generosa disposición y por la información proporcionada para la culminación de la investigación.



DEDICATORIA

A mis padres Gregorio y Juana, mis hermanos Heber y Wilder, quienes me han brindado su apoyo y amor incondicional para que pueda alcanzar mis metas.



A mi familia, en especial a mis padres Néstor, Benigna y hermanos, por su constante aliento, su amor inmenso y apoyarme siempre en todo lo que promuevo.

Sonia.

ÍNDICE

| | |
|---|----|
| ÍNDICE DE FIGURAS | 3 |
| ÍNDICE DE TABLAS | 7 |
| RESUMEN | 8 |
| ABSTRACT | 9 |
| INTRODUCCIÓN | 10 |
| CAPÍTULO 1: PLANTEAMIENTO DEL PROBLEMA DE INVESTIGACIÓN | 12 |
| 1.1. Descripción del Problema | 13 |
| 1.2. Justificación del Problema..... | 15 |
| 1.3. Objetivos de la Investigación..... | 15 |
| CAPÍTULO 2: MARCO TEÓRICO..... | 17 |
| 2.1. Antecedentes de la Investigación..... | 18 |
| 2.2. Sustento Teórico | 22 |
| 2.3. Glosario de Términos Básicos..... | 69 |
| 2.4. Hipótesis de la Investigación..... | 72 |
| 2.5. Operacionalización de Variables..... | 72 |
| CAPÍTULO 3: DISEÑO METODOLÓGICO DE INVESTIGACIÓN..... | 73 |
| 3.1. Tipo y Diseño de Investigación..... | 74 |
| 3.2. Población y Muestra de Investigación | 75 |
| 3.3. Ubicación y Descripción de la Población..... | 75 |
| 3.4. Material Experimental | 75 |
| 3.5. Técnicas e Instrumentos para Recolectar Información..... | 75 |
| 3.6. Técnicas para el Procesamiento y Análisis de Datos | 76 |
| 3.7. Procedimiento del Experimento..... | 76 |
| 3.8. Plan de Tratamiento de los Datos | 76 |
| 3.9. Diseño Estadístico para la Prueba de Hipótesis o Prueba de Hipótesis | 76 |
| CAPÍTULO 4: ANÁLISIS E INTERPRETACIÓN DE LOS RESULTADOS..... | 78 |
| 4.1. Comprensión de la Situación Actual..... | 79 |



4.2. Comprensión de los Datos85

4.3. Procesamiento de los Datos.....106

4.4. Evaluación del Modelo153

CONCLUSIONES 160

SUGERENCIAS 163

BIBLIOGRAFIA 165

ANEXOS 172



ÍNDICE DE FIGURAS

| | |
|---|----|
| Figura 2.1: Definición de Minería de Datos..... | 23 |
| Figura 2.2: Representación General de los Modelos de Minería de Datos. | 24 |
| Figura 2.3: Métodos del Modelo Descriptivo..... | 25 |
| Figura 2.4: Agrupamiento..... | 26 |
| Figura 2.5: Representación del Método de Correlación..... | 27 |
| Figura 2.6: Esquema General del Modelo Predictivo..... | 29 |
| Figura 2.7: Representación del Modelo Predictivo..... | 30 |
| Figura 2.8: Regresión Lineal Simple..... | 33 |
| Figura 2.9: Mean Absolute Error..... | 35 |
| Figura 2.10: Root Mean Squared Error..... | 37 |
| Figura 2.11: Relative Absolute Error..... | 38 |
| Figura 2.12: Fórmula T | 38 |
| Figura 2.13: Root Relative Squared Error..... | 39 |
| Figura 2.14: Fases de la Metodología CRISP-DM..... | 42 |
| Figura 2.15: Opciones de Trabajo de Weka..... | 58 |
| Figura 2.16: Preprocesado de los Datos..... | 59 |
| Figura 2.17: Estructura de la Toma de Decisiones..... | 60 |
| Figura 2.18: Número de Intervenciones Quirúrgicas..... | 61 |
| Figura 2.19: Número de Pacientes Operados..... | 61 |
| Figura 2.20: Aprovechamiento de Salas de Operación..... | 62 |
| Figura 2.21: Grado de Cumplimiento de la Operaciones Programadas..... | 63 |
| Figura 2.22: Rendimiento del Quirófano..... | 63 |
| Figura 2.23: Porcentaje de Intervenciones Quirúrgicas Suspendidas..... | 64 |
| Figura 2.24: Promedio de Días Preoperatorios..... | 65 |
| Figura 2.25: Porcentaje de Complicaciones Intraoperatorias..... | 65 |
| Figura 2.26: Porcentaje de Complicaciones Post Operatorias..... | 66 |
| Figura 2.27: Porcentaje de Complicaciones Anestésicas..... | 67 |
| Figura 2.28: Porcentaje de Complicaciones en Recuperación..... | 67 |
| Figura 2.29: Muerte o Defunción en el Centro Quirúrgico..... | 68 |
| Figura 3.1: Grupo Experimental..... | 74 |
| Figura 3.2: Fórmula de Precisión..... | 77 |
| Figura 3.3: Fórmula de Datos Históricos..... | 77 |
| Figura 3.4 Fórmula de Mejora en la Toma de Decisiones..... | 77 |
| Figura 4.1: Organigrama Departamento de Cirugía..... | 79 |
| Figura 4.2: Plan del Proyecto..... | 84 |

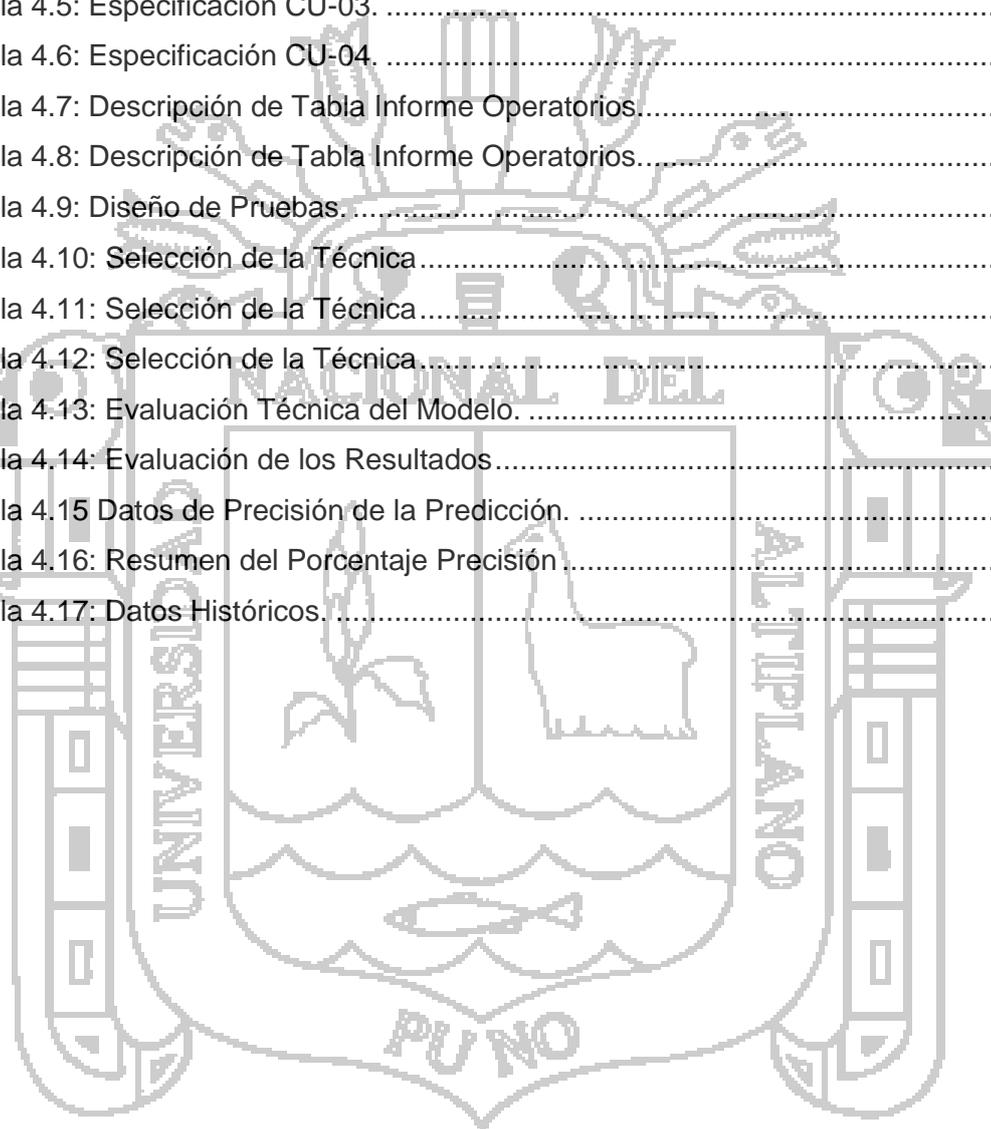
| | |
|--|-----|
| Figura 4.3: Arquitectura de la Herramienta para el Modelo..... | 88 |
| Figura 4.4: Caso de Uso Buscar Paciente (CU-01). | 89 |
| Figura 4.5: Caso de Uso Registrar Informe (CU-02)..... | 90 |
| Figura 4.6: Caso de Uso Modelar (CU-03). | 91 |
| Figura 4.7: Caso de Uso Ver Estadísticas. | 92 |
| Figura 4.8: Diagrama de Secuencia Buscar Paciente..... | 93 |
| Figura 4.9: Caso de Uso Registrar Informe | 94 |
| Figura 4.10: Diagrama de Secuencia Modelar..... | 95 |
| Figura 4.11: Diagrama de Secuencia Ver Estadística..... | 95 |
| Figura 4.12: Diagrama de Actividades Buscar Paciente..... | 96 |
| Figura 4.13: Diagrama de Actividades Buscar Paciente..... | 97 |
| Figura 4.14: Diagrama de Actividades Modelar..... | 97 |
| Figura 4.15: Diagrama de Actividades Ver Estadística..... | 98 |
| Figura 4.16: Diagrama de Clases..... | 99 |
| Figura 4.17: Interfaz de Buscar Paciente..... | 100 |
| Figura 4.18: Interfaz de Nuevo Paciente. | 100 |
| Figura 4.19: Interfaz de Modificar Informe Operatorio - A..... | 101 |
| Figura 4.20: Interfaz de Modificar Informe Operatorio – B..... | 101 |
| Figura 4.21: Modelo Entidad Relación..... | 102 |
| Figura 4.22: Exploración de datos..... | 105 |
| Figura 4.23: Consulta de Registros con Dato Nulo..... | 107 |
| Figura 4.24: Consulta NIQ..... | 108 |
| Figura 4.25: Consulta NPO. | 108 |
| Figura 4.26 Consulta ASO..... | 108 |
| Figura 4.27: Consulta GCOP..... | 109 |
| Figura 4.28: Consulta RQ..... | 109 |
| Figura 4.29: Consulta PIQS..... | 110 |
| Figura 4.30: Consulta PDP..... | 110 |
| Figura 4.31: Consulta PIPQ..... | 110 |
| Figura 4.32: Consulta PCI. | 111 |
| Figura 4.33: Consulta PCA..... | 111 |
| Figura 4.34: Consulta PCR..... | 111 |
| Figura 4.35: Consulta TM. | 112 |
| Figura 4.36: Formato del Filtro de la Instancia RemovePercentage. | 114 |
| Figura 4.37: Formato del algoritmo de Regresión Lineal | 115 |
| Figura 4.38: Modelo Regresión Lineal NIQ en Unid..... | 119 |
| Figura 4.39: Modelo Árboles de Decisión NIQ en Unid..... | 119 |

| | |
|---|-----|
| Figura 4.40: Predicción con Regresión Lineal NIQ | 120 |
| Figura 4.41: Predicción con Árboles de Decisión NIQ. | 120 |
| Figura 4.42: Evaluación con Regresión Lineal NIQ. | 121 |
| Figura 4.43: Evaluación con Árboles de Decisión NIQ. | 121 |
| Figura 4.44: Modelo con Regresión Lineal NPO en Unid..... | 122 |
| Figura 4.45: Modelo Árboles de Decisión NPO en Unid. | 122 |
| Figura 4.46: Predicción con Regresión Lineal NPO..... | 123 |
| Figura 4.47: Predicción con Árboles de Decisión NPO | 123 |
| Figura 4.48: Evaluación con Regresión Lineal NPO..... | 124 |
| Figura 4.49: Evaluación con Árboles de Decisión NPO. | 124 |
| Figura 4.50: Modelo con Regresión Lineal ASO en Horas..... | 125 |
| Figura 4.51: Modelo Con Árboles de Decisión ASO en Horas..... | 125 |
| Figura 4.52: Predicción con Regresión Lineal ASO..... | 126 |
| Figura 4.53: Predicción con Árboles de Decisión ASO. | 126 |
| Figura 4.54: Evaluación con Regresión Lineal ASO. | 127 |
| Figura 4.55: Evaluación con Árboles de Decisión ASO | 127 |
| Figura 4.56: Modelo con Regresión Lineal GCOP en Unid..... | 128 |
| Figura 4.57: Modelo con Árboles de Decisión GCOP en Unid..... | 128 |
| Figura 4.58: Predicción con Regresión Lineal GCOP..... | 129 |
| Figura 4.59: Predicción con Árboles de Decisión GCOP..... | 129 |
| Figura 4.60: Evaluación con Regresión Lineal GCOP..... | 130 |
| Figura 4.61: Evaluación con Árboles de Decisión GCOP..... | 130 |
| Figura 4.62: Modelo con Regresión Lineal RQ en Unid..... | 131 |
| Figura 4.63: Modelo con Árboles de Decisión RQ en Unid..... | 131 |
| Figura 4.64: Predicción con Regresión Lineal RQ..... | 131 |
| Figura 4.65: Predicción con Árboles de Decisión RQ..... | 132 |
| Figura 4.66: Evaluación con Regresión Lineal RQ..... | 132 |
| Figura 4.67: Evaluación con Árboles de Decisión RQ | 133 |
| Figura 4.68: Modelo con Regresión Lineal PIQS en Unid..... | 133 |
| Figura 4.69: Modelo con Árboles de Decisión PIQS en Unid..... | 133 |
| Figura 4.70: Predicción con Regresión Lineal PIQS..... | 134 |
| Figura 4.71: Predicción con Árboles de Decisión PIQS | 134 |
| Figura 4.72: Evaluación con Regresión Lineal PIQS..... | 135 |
| Figura 4.73: Evaluación con Árboles de Decisión PIQS..... | 135 |
| Figura 4.74: Modelo con Regresión Lineal PDP en días..... | 136 |
| Figura 4.75: Modelo con Árboles de Decisión PDP en días..... | 136 |
| Figura 4.76: Predicción con Regresión Lineal PDP..... | 137 |

| | |
|--|-----|
| Figura 4.77: Predicción con Árboles de Decisión PDP. | 137 |
| Figura 4.78: Evaluación con Regresión Lineal PDP. | 138 |
| Figura 4.79: Evaluación con Árboles de Decisión PDP. | 138 |
| Figura 4.80: Modelo con Regresión Lineal PIPQ en Unid..... | 139 |
| Figura 4.81: Modelo con Árboles de Decisión PIPQ en Unid..... | 139 |
| Figura 4.82: Predicción con Regresión Lineal PIPQ..... | 140 |
| Figura 4.83: Predicción con Árboles de Decisión PIPQ..... | 140 |
| Figura 4.84: Evaluación con Regresión Lineal PIPQ..... | 141 |
| Figura 4.85: Evaluación con Árboles de Decisión PIPQ..... | 141 |
| Figura 4.86: Modelo con Regresión Lineal PCI en Unid..... | 142 |
| Figura 4.87: Modelo con Árboles de Decisión PCI en Unid..... | 142 |
| Figura 4.88: Predicción con Regresión Lineal PCI..... | 143 |
| Figura 4.89: Predicción con Árboles de Decisión PCI..... | 143 |
| Figura 4.90: Evaluación con Regresión Lineal PCI..... | 144 |
| Figura 4.91: Evaluación con Árboles de Decisión PCI..... | 144 |
| Figura 4.92: Modelo con Regresión Lineal PCA en Unid..... | 145 |
| Figura 4.93: Modelo con Árboles de Decisión PCA en Unid..... | 145 |
| Figura 4.94: Predicción con Regresión Lineal PCA..... | 145 |
| Figura 4.95: Predicción con Árboles de Decisión PCA..... | 146 |
| Figura 4.96: Evaluación con Regresión Lineal PCA..... | 146 |
| Figura 4.97: Evaluación con Árboles de Decisión PCA..... | 147 |
| Figura 4.98: Modelo con Regresión Lineal PCR en porcentaje..... | 147 |
| Figura 4.99: Modelo con Árboles de Decisión PCR en porcentaje..... | 148 |
| Figura 4.100: Predicción con Regresión Lineal PCR..... | 148 |
| Figura 4.101: Predicción con Árboles de Decisión PCR..... | 149 |
| Figura 4.102: Evaluación con Regresión Lineal PCR..... | 149 |
| Figura 4.103: Evaluación con Árboles de Decisión PCR..... | 150 |
| Figura 4.104: Modelo con Regresión Lineal TM en Porcentaje..... | 150 |
| Figura 4.105: Modelo con Árboles de Decisión TM en Porcentaje..... | 150 |
| Figura 4.106: Predicción con Regresión Lineal TM..... | 151 |
| Figura 4.107: Predicción con Árboles de Decisión TM..... | 151 |
| Figura 4.108: Evaluación con Regresión Lineal TM..... | 152 |
| Figura 4.109: Evaluación con Árboles de Decisión TM..... | 152 |
| Figura 4.110: Fórmula Mejora en la Toma de Decisiones..... | 159 |

ÍNDICE DE TABLAS

| | |
|--|-----|
| Tabla 2.1: Operacionalización de Variables. | 72 |
| Tabla 4.1: Costos y Beneficios. | 82 |
| Tabla 4.2: Determinación de los Objetivos de Minería de Datos..... | 82 |
| Tabla 4.3: Especificación CU-01. | 89 |
| Tabla 4.4: Especificación CU-02. | 90 |
| Tabla 4.5: Especificación CU-03. | 91 |
| Tabla 4.6: Especificación CU-04. | 92 |
| Tabla 4.7: Descripción de Tabla Informe Operatorios..... | 103 |
| Tabla 4.8: Descripción de Tabla Informe Operatorios..... | 104 |
| Tabla 4.9: Diseño de Pruebas. | 114 |
| Tabla 4.10: Selección de la Técnica..... | 116 |
| Tabla 4.11: Selección de la Técnica..... | 117 |
| Tabla 4.12: Selección de la Técnica..... | 118 |
| Tabla 4.13: Evaluación Técnica del Modelo. | 153 |
| Tabla 4.14: Evaluación de los Resultados..... | 154 |
| Tabla 4.15 Datos de Precisión de la Predicción. | 156 |
| Tabla 4.16: Resumen del Porcentaje Precisión..... | 157 |
| Tabla 4.17: Datos Históricos..... | 158 |



RESUMEN

La presente investigación titulada “MODELO PREDICTIVO BASADO EN MINERÍA DE DATOS PARA LA MEJORA EN LA TOMA DE DECISIONES DEL DEPARTAMENTO DE CIRUGÍA DEL HOSPITAL REGIONAL MANUEL NÚÑEZ BUTRÓN” , tuvo por objetivo general desarrollar un Modelo Predictivo basado en Minería de Datos para la mejora en la toma de decisiones del Departamento Cirugía del Hospital Regional Manuel Núñez Butrón. El problema detectado es la carencia de información requerida para lograr una adecuada toma de decisiones, además de los procesos que involucran obtener dicha información. En esta investigación se pone como caso de estudio el departamento de cirugía, el cual además de no poder obtener una predicción de sus indicadores hospitalarios, tiene el gran problema de registrar sus datos escritos a mano y por lo tanto existen datos ilegibles. En esta investigación se determinó que el criterio de éxito es el control de los indicadores hospitalarios, los cuales son base para una adecuada toma de decisiones, para ello se aplicó la técnica de regresión lineal del método de regresión del modelo predictivo de minería de datos aplicando la metodología CRISP-DM, utilizando el software Weka y para la recolección de datos se desarrolló un sistema de información. Finalmente, se realizó la prueba de la hipótesis que se demuestra empíricamente, el desarrollo de un “Modelo predictivo basado en minería de datos para la mejora en la toma de decisiones es aceptable ya que mejora en un 91.97%”.

Palabras claves: Minería de datos, CRISP-DM, modelo predictivo, weka, precisión.

ABSTRACT

This research entitled "PREDICTIVE MODEL BASED DATA MINING FOR IMPROVED DECISION MAKING THE DEPARTMENT OF SURGERY OF REGIONAL HOSPITAL MANUEL NÚÑEZ BUTRÓN" overall goal was to develop a system based on data mining to improve the Predictive Model decisions of the Department of Surgery Regional Hospital Manuel Núñez Butrón. The problem identified is the lack of information required for proper decision-making, and the processes involved to obtain such information. In this research as a case study puts the department of surgery, which besides not being able to get a prediction of your hospital indicators have the great problem of handwritten record their data and therefore are unreadable data. In this investigation it was determined that the criterion for success is control of hospital indicators, which are the basis for proper decision-making, thus the technique of linear regression of regression method of predictive data mining model was applied using the CRISP-DM methodology using the Weka software and data collection information system was developed. Finally, hypothesis testing is demonstrated empirically performed, the development of a "based on data mining to improve decision making in Predictive model is acceptable as it improves a 91.97%".

Key words: Data Mining, CRISP-DM, predictive model, weka, precision.

INTRODUCCIÓN

En la presente investigación se desarrolló un modelo predictivo que mejora la toma de decisiones en el Departamento de Cirugía del Hospital Regional Manuel Núñez Butrón aplicando la metodología CRISP-DM utilizando la técnica de Regresión Lineal.

Internamente el hospital atiende varias especialidades divididas en departamentos en el que a continuación se mencionan los más importantes: Medicina, Pediatría, Cirugía, Emergencia, Enfermería, Gineco Obstetricia, Patología, Estadística. Cada uno de estos Departamentos con procedimientos internos de atención, son atendidos por el departamento de estadística e informática, para el procesamiento de sus atenciones internas, control de sus indicadores y gestión de su toma de decisión.

Con la minería de datos se busca procesar analizar y predecir datos para mejorar la toma de decisiones, como caso de estudio en esta tesis se estudia el Departamento de cirugía.

Esta tesis se encuentra constituida por cuatro capítulos, todos ellos relacionados de manera que haya una coherencia entre las distintas partes y sea fácil su comprensión para los lectores de la presente investigación.

CAPÍTULO I: Planteamiento y formulación del problema, justificación de la investigación, objetivos de la investigación, limitaciones de la investigación. Estos puntos sitúan al lector dentro de la problemática de la investigación; así como la hipótesis que será demostrada y la operacionalización de variables.

CAPÍTULO II: El marco teórico compuesto de tres partes: antecedentes de la investigación, donde se considera Internamente el hospital atiende varias especialidades divididas en departamentos: medicina, pediatría, cirugía, emergencia, enfermería, gineco obstetricia, patología, cada una de ellas con procedimientos internos de atención, todos estos departamentos son atendidos por el departamento de estadística e informática, para el procesamiento de sus atenciones internas, control de sus indicadores y gestión de sus tomas de decisión como departamento.

Los antecedentes investigativos que sirvieron de base; el marco teórico, donde reside toda la teoría necesaria para el mejor entendimiento del presente trabajo; y finalmente el marco conceptual, necesario para comprender los términos básicos más utilizados a lo largo de la investigación.

CAPÍTULO III: El método de investigación constituido por el trabajo pre-experimental realizado; así como también de la metodología de desarrollo; las herramientas y técnicas empleadas para la recolección de datos y posterior análisis estadístico empleadas durante la investigación.

CAPÍTULO IV: Exposición y análisis de los resultados constituido por el desarrollo de la metodología CRISP-DM comprensión del negocio, comprensión de datos, preparación de datos, modelado, evaluación.

Conclusiones a las que se ha llegado como resultado de la implementación del modelo predictivo basado en minería de datos.

Recomendaciones y sugerencias para seguir investigando e implementando trabajos relacionados.



CAPÍTULO 1: PLANTEAMIENTO DEL PROBLEMA DE INVESTIGACIÓN

1.1. Descripción del Problema

Los altos niveles competitivos que deben alcanzarse en las organizaciones de todos los sectores, obliga a sus responsables a apoyarse para la toma de decisiones en información directa, real, veraz y demostrable que evidencie la gestión realizada por cada uno de los miembros de la institución (MINSA, 2013).

Los indicadores de gestión en el sector salud, son herramientas que permiten medir la administración que buscan optimizar la oferta hospitalaria a una demanda de necesidades de atención de salud que les permita tomar decisiones para la gestión administrativa eficiente y oportuna (MINSA, 2013), bajo este esquema el Hospital Regional Manuel Núñez Butrón de Puno que recibe una gran demanda de pacientes en el 2014 se registraron en promedio 28422 casos atendidos (Hospital Manuel Núñez Butrón, 2013).

Para la adecuada gestión hospitalaria es necesario que internamente cada departamento: medicina, pediatría, cirugía, emergencia, enfermería, gineco obstetricia, patología, tenga un procesamiento de datos, control de sus indicadores y estos sirvan de apoyo en la gestión de tomas de decisión a nivel.

Como caso de estudio se tiene el departamento de cirugía que presenta grandes problemas con la información requerida para el procesamiento, análisis y predicción de datos para toma de decisiones ya que mensualmente es necesario reportar los siguientes indicadores (MINSA, 2013):

- Número de intervenciones quirúrgicas.
- Número de pacientes operados.
- Aprovechamiento de salas de operaciones.
- Grado de cumplimiento de las operaciones programadas.

- Rendimiento quirófano.
- Porcentaje de intervenciones quirúrgicas suspendidas.
- Promedio de días preoperatorios.
- Porcentaje de infecciones postquirúrgicas.
- Porcentaje de complicaciones intraoperatorias.
- Porcentaje de complicaciones anestésicas.
- Porcentaje de complicaciones en recuperación.
- Tasa de mortalidad en el centro quirúrgico.

De la información requerida, el departamento de cirugía solo puede reportar 6 indicadores (Número de Pacientes Operados, Número de Intervenciones Quirúrgicas, Aprovechamiento de Salas de operación, Rendimiento del Quirófano, Porcentaje de Intervenciones Quirúrgicas Suspendidas, Tasa de Mortalidad) (Hospital Regional Manuel Núñez Butrón, 2014), ya que el procesamiento de datos resulta dificultoso.

De seguir con esta situación, el Departamento de Cirugía seguirá registrando en formatos denominados “informe operatorio” donde los datos son ilegibles lo que dificulta el análisis y predicción de los mismos en consecuencia no colabora en la toma de decisiones.

1.2. Formulación del Problema

¿En qué medida el Modelo Predictivo Basado en Minería de Datos mejora la toma de decisiones del Departamento de Cirugía del Hospital Regional Manuel Núñez Butrón?.

1.3. Justificación del Problema

El personal directivo del hospital que gestiona la toma de decisiones no tiene la información adecuada para gestionar los recursos internos del hospital, con la investigación se les brinda datos del departamento de cirugía para optimizar la administración de los servicios a nivel operativo, utilizando como medio la minería de datos la cual se fundamenta en la busca de patrones dentro de grandes bases de datos utilizando métodos predictivos, haciendo uso de recursos informáticos y tecnológicos.

Con la investigación se aprovecha los datos que se registra en cada una de las atenciones en el Departamento de Cirugía con el fin de obtener un porcentaje de precisión clara de todos los indicadores que se deben reportar como departamento mensualmente para contribuir con una adecuada toma de decisiones.

El desarrollo del modelo predictivo basado en minería de datos permite analizar, procesar y predecir los indicadores que sirven para la toma de decisiones ya que con ello se tiene el acceso inmediato a los datos.

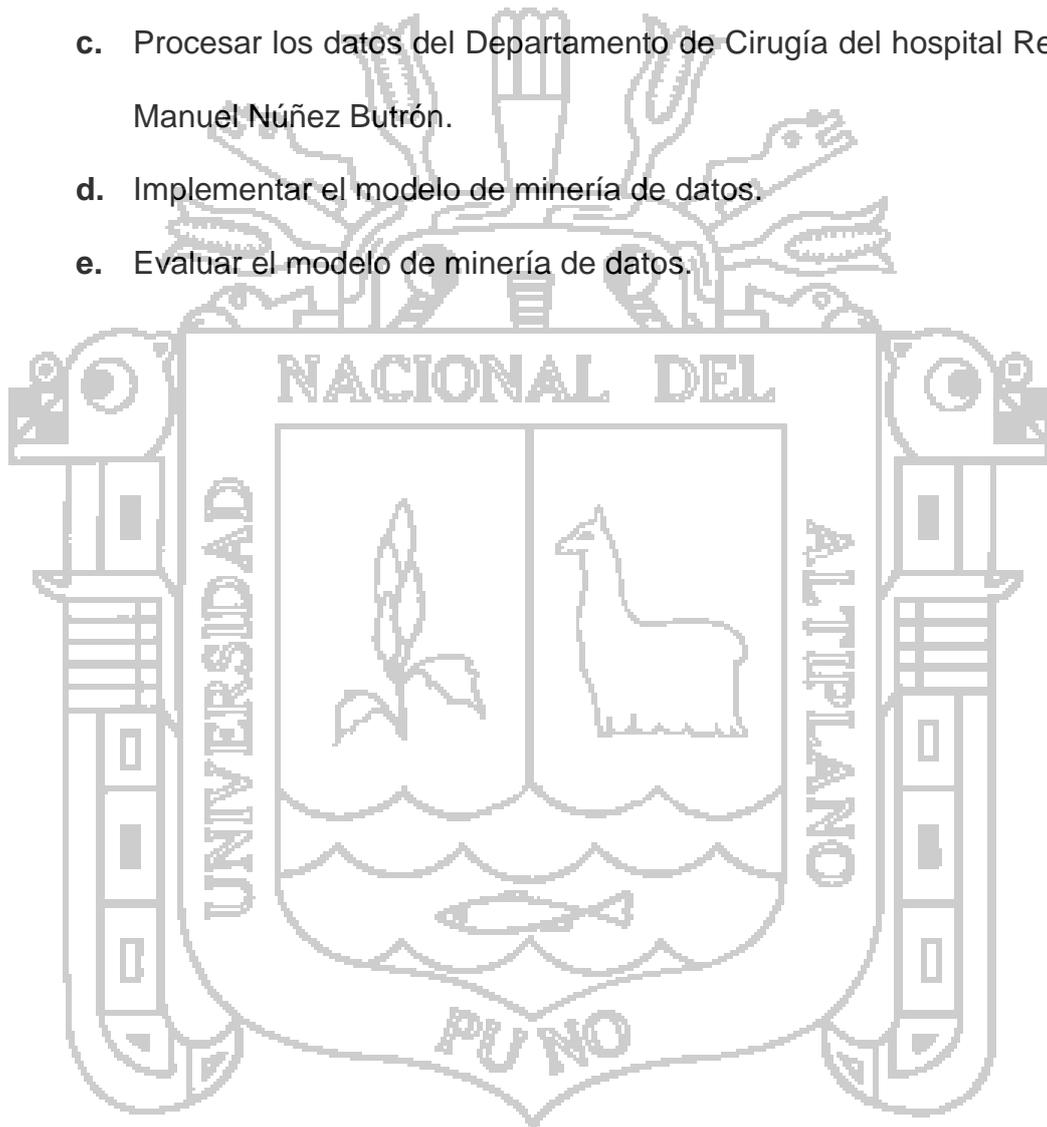
1.4. Objetivos de la Investigación

1.4.1. Objetivo General

Desarrollar un Modelo Predictivo Basado en Minería de Datos para la Mejora en la Toma de Decisiones del Departamento Cirugía del Hospital Regional Manuel Núñez Butrón.

1.4.2. Objetivos Específicos

- a. Comprender la situación actual del Departamento de Cirugía del hospital Regional Manuel Núñez Butrón.
- b. Comprender los datos del Departamento de Cirugía del hospital Regional Manuel Núñez Butrón.
- c. Procesar los datos del Departamento de Cirugía del hospital Regional Manuel Núñez Butrón.
- d. Implementar el modelo de minería de datos.
- e. Evaluar el modelo de minería de datos.





2.1. Antecedentes de la Investigación

2.1.1. “Portal Web de Noticias del Perú Aplicando Técnicas de Recuperación de Información y Minería De Datos de Alain Melquiades Alejo Huarachi de la Universidad Nacional del Altiplano Puno”.

En el trabajo de investigación mencionado Alain Alejo Huarachi (Alejo Huarachi, 2009), llega a las siguientes conclusiones:

Primera: El portal web de noticias del Perú, que aplica técnicas de recuperación de información y minería de datos, logra extraer, clasificar y visualizar las noticias publicadas en otros sitios web de diversos medios de comunicación del Perú.

Segunda: El análisis permitió identificar cuatro casos de uso principales: buscar, ver, extraer y clasificar noticia; cinco entidades de la base de datos: fuente, categoría, noticia, etiqueta y comentario, seis diagramas de secuencia y dos diagramas de actividades; así como la infraestructura para el funcionamiento del portal web.

Tercera: El diseño permitió separar el sistema en dos módulos: presentación y recuperación, así como determinar los mecanismos de navegación y la interfaz de usuario para tres tipos de búsqueda: por departamento, por categoría y por términos; y para ver una noticia en detalle.

2.1.2. “Desarrollo de un Modelo de Minería de Datos Para la Toma de Decisiones en la Gestión de Inventarios en Empresas Comerciales de Sandra Jaqueline Mendoza Castillo y Fernando Zavaleta Henríquez de la Universidad Nacional de Trujillo”.

En el trabajo de Investigación mencionado Sandra Jaqueline Mendoza Castillo y Fernando Zavaleta Henríquez (Sandra Jaqueline Mendoza Castillo, Fernando Zavaleta Henríquez, 2013), llegan a las siguientes conclusiones:

De la investigación bibliográfica se determinó que la minería de datos es el conjunto de herramientas y técnicas de análisis de datos que permiten crear escenarios, de los cuales se puede obtener información útil para la toma de decisiones.

Del análisis comparativo de las técnicas de minería de datos se pudo observar que la técnica de redes neuronales presenta la mayoría de los criterios, sin embargo son muy costosas computacionalmente por lo que se decidió optar por los árboles de decisión apoyados de regresión lineal múltiple.

Para el diseño del modelo se propuso una arquitectura por niveles de manera que haya una separación entre los distintos procesos que involucra la minería de datos y de forma que la mayoría de la interacción ocurra únicamente entre niveles vecinos.

La implementación del modelo propuesto es un objetivo que no se pudo alcanzar, sin embargo, para llevarlo a cabo será necesario un análisis

minucioso sobre los datos recolectados de la empresa elegida para el caso de estudio.

2.1.3. “Aplicación de Técnicas de Minería de Datos a Datos Obtenidos por el Centro Andaluz de Medio Ambiente (CEAMA) de Francisco José García Gonzales de la Universidad de Granada”.

En el trabajo de investigación mencionado Francisco José García Gonzales (Gonzales F. J., 2013), llega a las siguientes conclusiones:

Tanto el clasificador Naive Bayes, como el Stacking y el J48 ofrecen un porcentaje de instancias correctamente clasificadas superior al 81%, si bien, como se ha comprobado con el clasificador OneR existe un atributo (POR_3) que permite clasificar de forma correcta entorno a un 64% de las instancias.

En relación a los diferentes modos de prueba, es el basado en la muestra de entrenamiento el que ofrece mejores resultados para cada uno de los clasificadores. Esto es lógico puesto que este clasificador se evalúa en el mismo conjunto sobre el que se creó el modelo de clasificación produciendo una sobrestimación de los resultados. Sin embargo esta situación no es ratificada por el clasificador OneR donde el porcentaje de instancias bien clasificadas mediante una división del fichero (64, 6004%), es superior al obtenido mediante el conjunto de entrenamiento (63,5483%).

En relación al tipo de clasificador, es el meta-clasificador Stacking el que mejores resultados ofrece, concretamente los valores de los índices Kappa para cada modo de prueba (0,8077, 0,7664 y 0,7714) son

superiores a sus correspondientes modos de prueba de los tres clasificadores.

La extracción de modelos predictivos mediante la selección evolutiva estratificada presenta el mejor comportamiento de entre los algoritmos de selección de instancias analizados. Ofrecen los árboles de decisión más pequeños con índices de precisión altos, similares a los que muestra el C4.5 sin reducción. La selección evolutiva estratificada permite obtener modelos predictivos con el equilibrio adecuado entre interoperabilidad y precisión.

2.1.4. “Minería de Datos con Weka para la Predicción del Precio de Automóviles de Segunda Mano de Francisco José García Gonzales de la Universidad Politécnica de Valencia”.

En el trabajo de investigación mencionado Agustín José Calleja Gómez (Gomez, 2010), llega a las siguientes conclusiones:

Como hemos podido ver y hemos explicado durante la memoria de este proyecto, la minería de datos es una herramienta con un potencial increíble y aplicable en un sinfín de proyectos, circunstancias y finalidades.

A través de este proyecto, nos hemos podido introducir un poco en este mundo de la obtención de información relevante de masas de datos compactas. Hemos introducido el concepto de minería de datos, así como analizado varios de sus numerosos tipos y familias de métodos de cálculo y obtención de resultados.

Evidentemente, se puede profundizar muchísimo más en la minería de datos, pero se escapa del objetivo de este proyecto. No obstante, hemos podido aprender a utilizar una herramienta muy potente y con licencia OpenGPL para la minería de datos llamada Weka, hemos analizado sus características y hemos surcado por encima de sus numerosas posibilidades y funcionalidades.

2.2. Sustento Teórico

2.2.1. Minería de Datos

Se encontraron varias definiciones de Minería de Datos. A Continuación se mencionan algunas:

Gabriel Amorocho define que: “Es el proceso de recolectar y analizar grandes cantidades de datos (típicamente históricos) desde diferentes perspectivas para identificar patrones, correlaciones o tendencias ocultas entre muchas variables para posteriormente inferir las razones subyacentes para estos comportamientos y proyectarlos al futuro” (Amorocho, 2015).

Otra definición de Palomino Mariana Garrido consiste en: “La aplicación de técnicas en grandes volúmenes de datos para descubrir información útil, aplicable y no trivial”. Esta definición, aplicada a un entorno más empresarial podría reconstruirse como “el conjunto de métodos, que junto con un profundo conocimiento del negocio, están orientados a identificar, en grandes volúmenes de datos, relaciones y tendencias ocultas hasta el momento” (Garrido, 2012).

Carlos Márquez Vera compara la minería de datos con la estadística: “La minería de datos está orientada a trabajar con cantidades muy grandes de datos (millones y billones de datos). En cambio la estadística no suele funcionar tan bien en bases de datos de tan gran tamaño y alta dimensionalidad” (Carlos Marquéz Vera, 2012).

Luis Carlos Molina Felix resalta que minería de datos: “Es un conjunto de áreas (Estadística, inteligencia artificial, reconocimiento de patrones, computación gráfica y bases de datos) que tienen como propósito la identificación de conocimiento obtenido a partir de las bases de datos que aporten un sesgo o apoyo hacia la toma de decisión” (Molina Felix, 2014).

De las definiciones anteriores concluimos que la minería de datos es el procesamiento de grandes cantidades de datos generando como resultado el conocimiento, ver Figura 2.1.

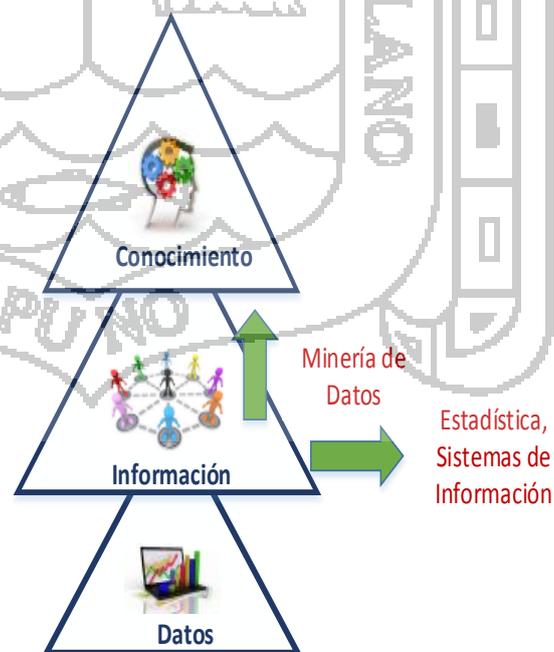


Figura 2.1: Definición de Minería de Datos.

2.2.2. Modelos de Minería de Datos

En la práctica, los modelos de minería de datos pueden ser de dos tipos: predictivos y descriptivos, otros autores los dividen en técnicas de minería de datos. Santiago Zapata menciona que: “se clasifican en dos grandes categorías: supervisadas o predictivas y no supervisadas o descriptivas” (Zapata, 2011). En la Figura 2.2 se muestra la representación de los modelos de Minería de Datos:

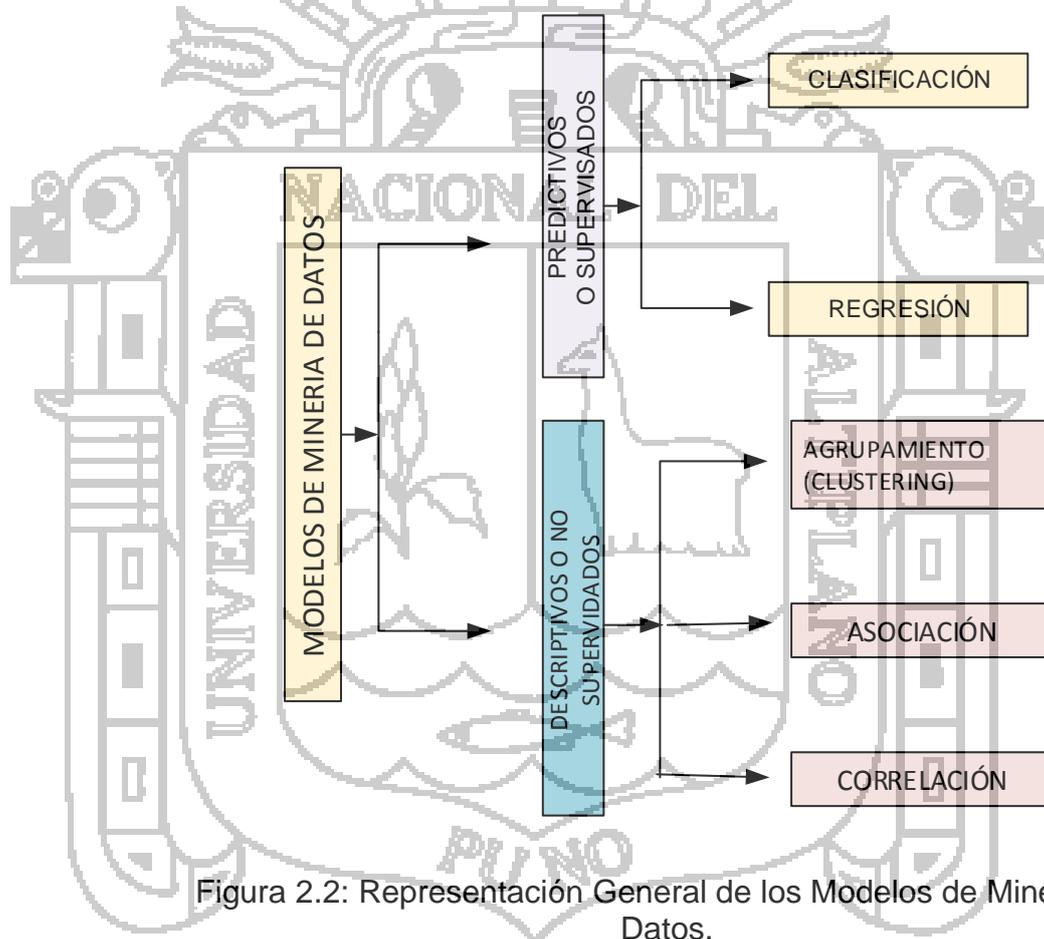


Figura 2.2: Representación General de los Modelos de Minería de Datos.

2.2.3. Modelo Descriptivo

Waldo Hasperué resalta que: “El modelo descriptivo identifica patrones que explican o resumen los datos, es decir, que sirve para explorar las propiedades de los datos examinados (Hasperué, Extracción de

Conocimiento en Grandes Bases de Datos Utilizando Estrategias Adaptativas, 2012).” Zapata, Santiago rescata que: “Los modelos descriptivos se clasifican en los métodos de: agrupamiento, asociación y correlación como se muestra en la Figura 2.3. Los modelos descriptivos también son llamados aprendizajes no supervisados ya que los modelos actúan solos y tratan de encontrar cosas en los datos” (Zapata, Técnicas de Minería de Datos basadas en Aprendizaje Automático , 2011).

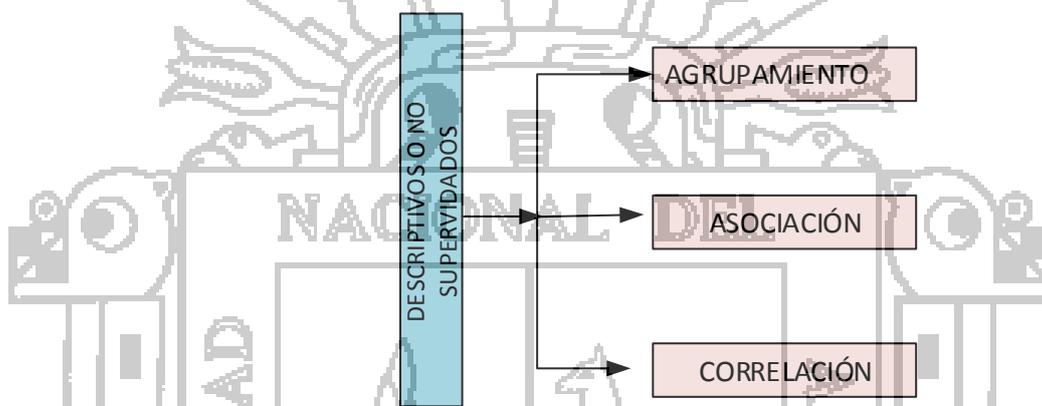


Figura 2.3: Métodos del Modelo Descriptivo.

a. Agrupamiento:

Daniela Godoy resalta agrupamiento: “Como un método no supervisado para dividir ejemplos en grupos cuando no existen categorías predefinidas. El aprendizaje no supervisado es un método descriptivo para interpretar un conjunto de datos, es el proceso de buscar un agrupamiento natural en un conjunto de datos en base a su similitud. Ver Figura 2.4.

El objetivo de agrupamiento es encontrar agrupaciones que estén lo más cercanos posibles malo es que se trabajan todos contra todos” (Godoy, 2015).

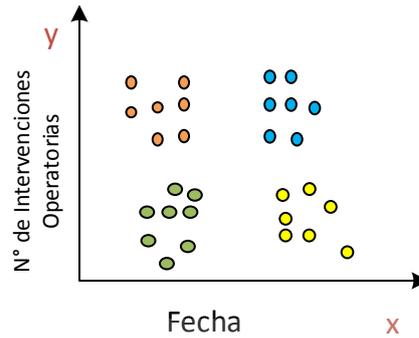


Figura 2.4: Agrupamiento.

b. Asociación

Según Haydeé Gómezz y María de los Angeles Cerón resaltan que: “El método de asociación detecta eventos que ocurren de manera simultánea” (Gómez Díaz, Haydeé; Cerón Reyes, María de los Angeles, 2010).

Jorge Antonio Párraga Álava y Robert Abel Alcívar Cevallos rescatan que: “El objetivo de las reglas de asociación que es encontrar asociaciones o correlaciones sobre los elementos u objetos de bases de datos transaccionales, relacionales o datawarehouse” (Párraga Álava, Jorge Antonio ; Alcívar Cevallos, Robert Abel, 2012).

c. Correlación

José Hernández, María Ramírez y Cesar Ferri definen: “Las correlaciones como una tarea descriptiva que se usan para determinar el grado de similitud de los valores de dos variables numéricas. Un mecanismo estándar para medir la correlación, para un caso llamado R, el cual es un valor real comprendido entre -1 y 1. Así, si R es 1 las variables están totalmente correlacionadas; si R es -1 las variables están totalmente

relacionadas negativamente; y si R es cero no existe correlación. Por consiguiente, cuando R es positivo, las variables tienen un comportamiento similar y cuando R es negativo una variable crece y la otra decrece” (Hernandez Orallo, José; Ramirez Quintana, María José; Ferri Ramirez, Cesar;, 2004).

José Hernández, María Ramírez y Cesar Ferri resaltan que: “una manera de visualizar la posible correlación entre las observaciones de dos variables (X e Y), es a través de un diagrama de dispersión como se muestra en la Figura 2.5 en el cual los valores que toman estas variables son representados por puntos. Su principal desventaja es que no puede ser utilizada para hacer predicciones, ya que no es clara la forma que toma la relación” (Hernandez Orallo, José; Ramirez Quintana, María José; Ferri Ramirez, Cesar;, 2004).

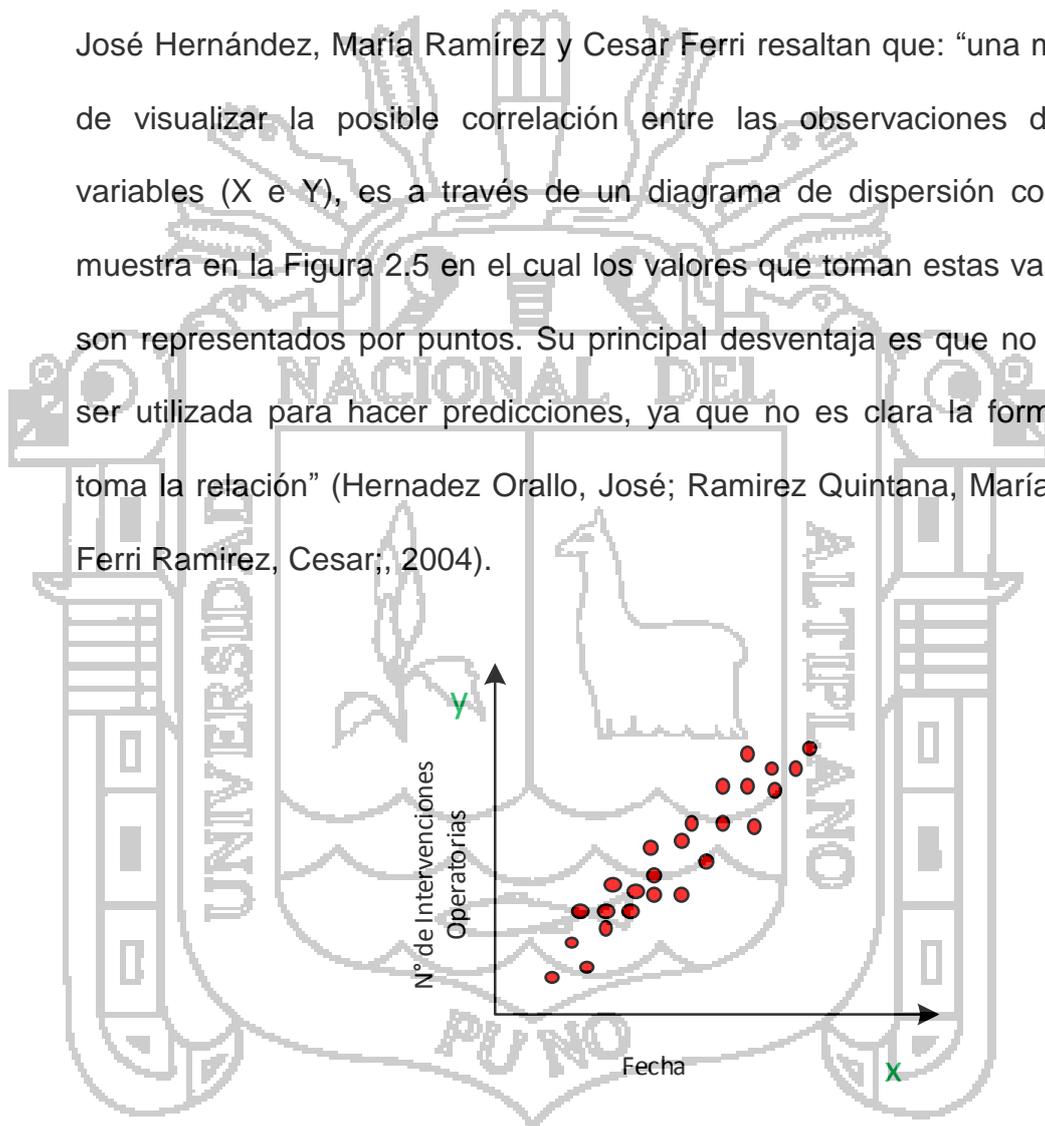


Figura 2.5: Representación del Método de Correlación.

2.2.4. Modelo Predictivo

Oldemar Rodríguez Rojas define modelo predictivo como: “entrenar modelos que se supervisa se dirige a algo para detectar algo o predecir.

El modelo predictivo es también conocido como Aprendizaje Supervisado, Método Predictivo, Estadistical learning, Analitics” (Rodríguez Rojas, Metodos Predictivos en Minerría de Datos - PERSON, 2015).

Sandra Mendoza y Fernando Henríquez resaltan que: “el modelo predictivo se emplea para estimar valores futuros de variables de interés. El proceso se basa en la información histórica de los datos mediante las cuales se predice un comportamiento de los datos, ya sea mediante clasificaciones, categorizaciones o regresiones. Los modelos predictivos siguen un aprendizaje supervisado, que consiste en aprender mediante el control de un supervisor o maestro que determina la respuesta que se desea generar del sistema. El atributo a predecir se conoce como variable dependiente u objetivo, mientras que los atributos utilizados para realizar la predicción se llaman variables independientes o de exploración” (Mendoza Castillo, Sandra Jaqueline; Zavaleta Henríquez, Fernando, 2013).

José Mondragón resalta que los modelos predictivos: “son parte fundamental de la minería de datos y por ello siempre es necesario revisar su desempeño y poder de predicción para ser mejorados o inclusive cambiados en sus parámetros ya que los patrones de comportamiento pueden variar según la época o momento en el que fueron desarrollados” (Mondragón, 2012).

Álvaro Jiménez y Hugo Álvarez destacan que: “en predicción, el objetivo es desarrollar un modelo que pueda inferir una variable a partir de alguna

combinación de otras variables incluidas en los datos. La predicción requiere etiquetas para la variable de salida para un conjunto de datos limitado, donde una etiqueta suponga una información fiable sobre el valor de la variable de salida en casos específicos. De todas maneras, en algunos casos es importante considerar el grado en el que estas etiquetas puedan ser aproximadas o inciertas. En un segundo tipo de uso, los métodos de predicción son utilizados para predecir cuál será el valor de salida en contextos donde no es deseable obtener una etiqueta para una construcción (por ejemplo, en ocasiones en las que no haya datos etiquetados)”. (Jímenez Galindo, Álvaro; Álvarez García, Hugo, 2010). En la Figura 2.6 se muestra el esquema general de modelo predictivo.

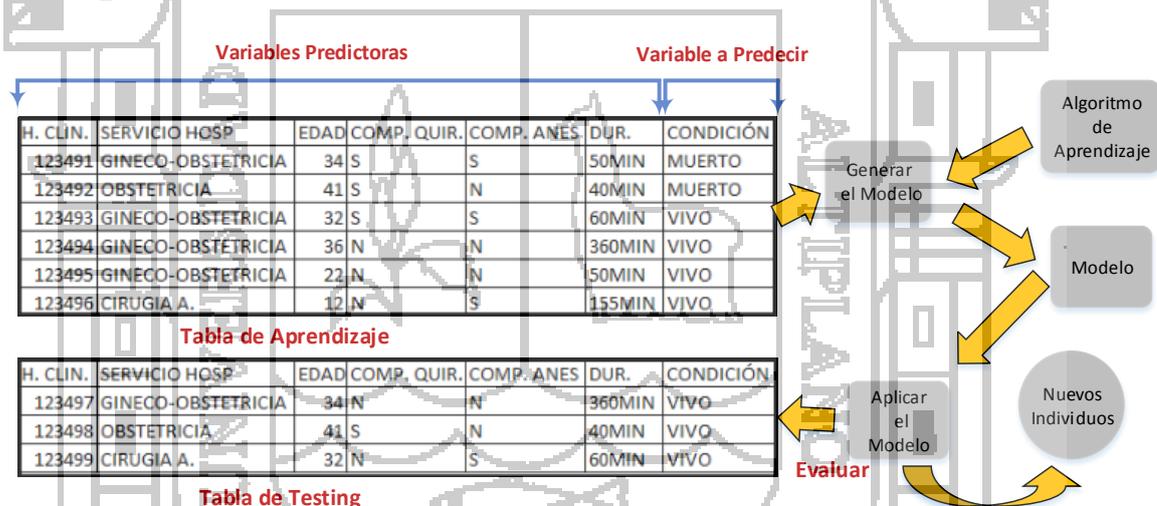


Figura 2.6: Esquema General del Modelo Predictivo.

2.2.5. Métodos del Modelo Predictivo

Según Gloria Hernández Gómez: “los modelos predictivos se pueden clasificar en dos métodos: Clasificación y Regresión” como se muestra en la Figura 2.7 representación del modelo predictivo (Hernández Gómez, 2014).

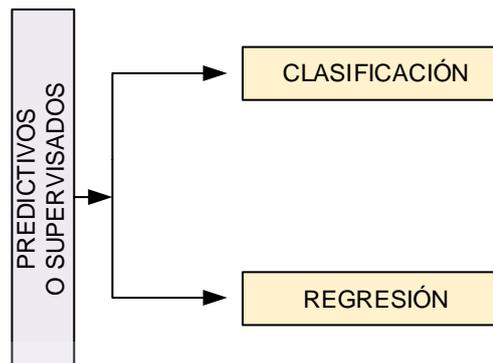


Figura 2.7: Representación del Modelo Predictivo.

- **Clasificación:**

Oldemar Rodríguez Rojas resalta que: “el método de clasificación se representa de la siguiente manera que: dada una colección de registros o conjunto de entrenamiento cada registro contiene un conjunto de variables que son los atributos denominadas x también llamadas variables predictoras, con una variable adicional que es la clase denominada y o variable a predecir” (Rodríguez Rojas, Metodos Predictivos en Minerría de Datos - PERSON, 2015).

Oldemar Rodríguez Rojas también destaca que: “el objetivo de la Clasificación que es encontrar un modelo (una función o algoritmo) para predecir la clase a la que pertenecía cada registro, esta asignación de una clase se debe hacer con la mayor precisión posible. Un conjunto de prueba (tabla de testing) se utiliza para determinar la precisión del modelo. Por lo general, el conjunto de datos dado se divide en dos conjuntos al azar: el de entrenamiento y el de prueba” (Rodríguez Rojas, Metodos Predictivos en Minerría de Datos - PERSON, 2015).

Oldemar Rodríguez define clasificación: “en una forma matemática dada una base de datos $D = \{t_1, t_2, t_3, \dots, t_n\}$ de tuplas o registros (individuos) y un conjunto de clases $C = \{C_1, C_2, C, \dots, C_m\}$, el problema de la clasificación es encontrar una función $f: D \rightarrow C$ tal que cada t_i es asignada una clase C_j . La función $f: D \rightarrow C$ podría ser una Red Neuronal, una Regresión Lineal, un árbol de Decisión y otros” (Rodríguez Rojas, Metodos Predictivos en Minería de Datos - PERSON, 2015).

- **Regresión:**

Jay Devore rescata que: “el término regresión fue utilizado por primera vez como un concepto estadístico en 1877 por Sir Francis Galton, quien llevó a cabo un estudio que mostró que la estatura de los niños nacidos de padres altos tiende a retroceder o “regresar” hacia la estatura media de la población. Designó la palabra regresión como el nombre del proceso general de predecir una variable (la estatura de los niños) a partir de otra (la estatura del padre o de la madre). Más tarde, los estadísticos acuñaron el término regresión múltiple para describir el proceso mediante el cual se utilizan varias variables para predecir otra” (Devore, 2008).

Héctor Rodríguez, destaca que: “la Regresión ayuda a descubrir la dependencia del valor de un atributo con respecto a otros atributos dentro de la misma entidad u objeto. La regresión es similar a los árboles de decisión en cuanto a su contribución para clasificar datos, pero predice atributos continuos, en lugar de separados” (Rodríguez, 2009).

En muchas investigaciones estadísticas de minería de datos tendientes a la toma de decisiones de tipo profesional o personal uno de los objetivos

principales es establecer relaciones que permitan pronósticar una o más variables en términos de otras.

2.2.6. Técnicas del Modelo Predictivo

Wilson Andrés Castillo Rojas Resalta que: “las siguientes Técnicas son utilizados en el Modelo Predictivo” (Castillo Rojas, 2013):

- Redes Neuronales.
- Árboles de Decisión (ID3,C4.5.C5.0).
- Árboles de Decisión (CART).
- Otros Árboles de Decisión.
- Regresión Lineal y Logarítmica.
- Regresión Logística.
- Naive Bayes.
- Vecinos Más Próximos.
- Algoritmos Genéticos y Evolutivos.
- Máquinas de Vectores de Soporte.
- CN2 Reglas.
- Análisis Discriminante Multivariante.

2.2.7. Regresión Lineal

Jay Devore destaca que: “el modelo de pronóstico de regresión lineal permite hallar el valor esperado de una variable aleatoria a cuando b toma un valor específico. La aplicación de este método implica un supuesto de linealidad cuando la demanda presenta un comportamiento creciente o decreciente, por tal razón, se hace indispensable que previo a la selección

de este método exista un análisis de regresión que determine la intensidad de las relaciones entre las variables que componen el modelo” (Devore, 2008).

Tipos de Modelos de Regresión Lineal

Los tipos de regresión lineal se clasifican de acuerdo a sus parámetros:

- **Regresión Lineal Simple:**

Irene Peláez resalta que: “la regresión lineal simple se basa en estudiar los cambios en una variable, no aleatoria, afectan a una variable aleatoria, en el caso de existir una relación funcional entre ambas variables que puede ser establecida por una expresión lineal, es decir, su representación gráfica es una línea recta como se muestra en la Figura 2.8. Es decir, se está en presencia de una regresión lineal simple cuando una variable independiente ejerce influencia sobre otra variable dependiente ejemplo: $Y = f(x)$ ” (Moral Peláez, 2012).

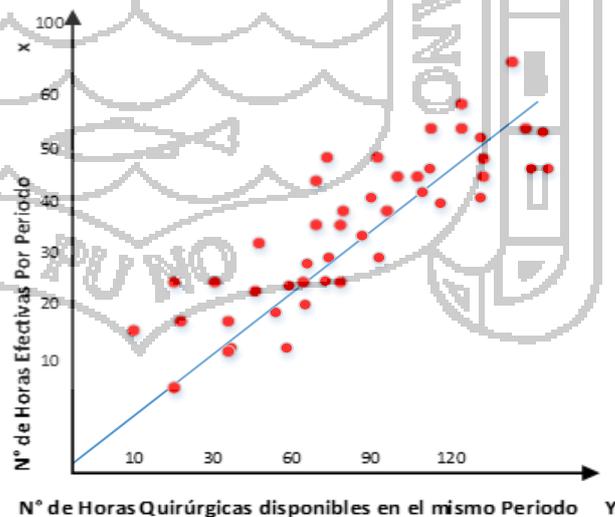


Figura 2.8: Regresión Lineal Simple.

- **Regresión Lineal Múltiple:**

José Manuel Rojo prevalece que: “la regresión lineal permite trabajar con una variable a nivel de intervalo o razón, así también se puede comprender la relación de dos o más variables y permitirá relacionar mediante ecuaciones, una variable en relación a otras variables llamándose Regresión múltiple. La regresión lineal múltiple es cuando dos o más variables independientes influyen sobre una variable dependiente ejemplo: $Y = f(x, w, z)$ ” (Rojo Abuín, 2017).

2.2.8. Parámetros para Calcular el Error del Modelo en Regresión Lineal

Paula Andrea Vizcaino Garzon resalta que: “los parámetros que utiliza weka para calcular el error del modelo aplicando la técnica de regresión lineal son los siguientes” (Vizcaino Garzon, 2008):

a. Coeficiente de Correlación (Correlation Coefficient):

Varey rescata que el coeficiente de correlación: “es el índice que mide la relación lineal entre dos variables aleatorias cuantitativas. A diferencia de la covarianza la correlación es independiente de la escala de medida de las variables. El coeficiente de correlación, comúnmente identificado como r o R , es una medida de asociación entre las variables aleatorias X y Y , cuyo valor varía entre -1 y $+1$ ” (Vararey, 2015) ver la siguiente Figura 2.9

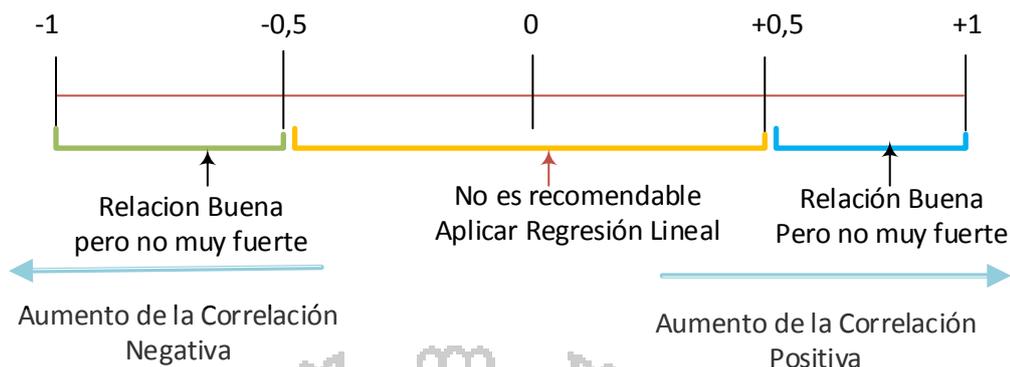


Figura 2.9: Coeficiente de Correlación.

b. Error Medio Absoluto (Mean Absolute Error):

Torre de la Vega rescata que: “error absoluto de una medida la diferencia entre el valor medio obtenido y el hallado en esa medida todo en valor absoluto. El error absoluto nos indica el grado de aproximación y da un indicio de la calidad de la medida. Indica la media del error producido en cada predicción. Mide la magnitud media de los errores en un conjunto de cálculos, sin tener en cuenta su dirección. Esto da la medida de precisión para las variables continuas. En otras palabras, el error medio absoluto es la media de la muestra de verificación de los valores absolutos de las diferencias entre los cálculos y la correspondiente observación. Es un resultado lineal, lo que significa que todas las diferencias individuales se ponderan por igual a la media. La función está dada por:

$$MAE = \frac{1}{n} \sum_{i=1}^n |f_i - y_i| = \frac{1}{n} \sum_{i=1}^n |e_i|$$

Figura 2.9: Mean Absolute Error.

Dónde:

f_i : Es la predicción

y_i : Es el valor verdadero.

MAE: Error medio absoluto.

Entonces el promedio de error absoluto, es la suma de los errores absolutos de clasificación en cada uno de los sujetos llevados a promedio. El clasificador que arroje mayor cifra (mayor a 0.1) define un error de clasificación alto, por lo cual no se debe considerar por sobre los que arrojen una cifra menor” (Torre de la Vega, 2012).

c. Error Cuadrático Medio (Root Mean Squared Error):

Pilar Gonzales define error cuadrático medio como: “una medida de uso frecuente de las diferencias entre los valores pronosticados por un modelo o un estimador y los valores realmente observados. El error cuadrático medio es una buena medida de precisión, pero sólo para comparar diferentes errores de predicción dentro de un conjunto de datos y no entre los diferentes, ya que es dependiente de la una escala muestra. Estas diferencias individuales también se denominan residuos, y el error cuadrático medio sirve para agregarlos en una sola medida de la capacidad de predicción. Es raíz cuadrada de la suma de los cuadrados de los errores individuales de las lecturas, entendiendo por tales a sus diferencias respecto del valor medio medido, que se adopta como valor verdadero convencional” (Gonzales P. , 2010).

David Freedman resalta que error cuadrático medio es: “una regla que mide la magnitud media del error. Esto es, la diferencia entre lo pronosticado y los correspondientes valores observados al cuadrado para que luego sea promediado a lo largo de la muestra. Por último, se toma la raíz cuadrada de la media. Dado que los errores son al cuadrado antes de que se promediaran, el error cuadrático medio da un peso relativamente alto a los grandes errores. Esto significa que el error cuadrático medio es más útil cuando los grandes errores son particularmente indeseables. Se debe tener en cuenta que el error cuadrático medio será siempre mayor o igual al error medio absoluto, la gran diferencia entre ellos, es en los errores individuales de la muestra. Si el $RMSE = MAE$, entonces todos los errores son de la misma magnitud. Ambos pueden ir de 0 a ∞ y son orientados a que los valores más bajos son los mejores:

$$E_i = \sqrt{\frac{1}{n} \sum_{j=1}^n (P_{(ij)} - T_j)^2}$$

Figura 2.10: Root Mean Squared Error.

Dónde:

$P_{(ij)}$: Es el valor que se ha predicho individualmente i del caso j , y T_j es el valor objetivo para el caso j . Es así que $P_{(ij)} = T_j$ y E_i son los rangos de los índices de 0 a infinito, donde 0 corresponde al ideal” (Freedman, 2009).

d. Error Relativo Absoluto (Relative Absolute Error):

La Universidad Autónoma del Estado de México prevalece que el error relativo absoluto: “es el error relativo a cada característica de la clase. El error relativo absoluto es el cociente entre el error absoluto y el que se da como representativo (la media aritmética). Es aquel que ayuda a predecir un valor relativo, que no es más que la media de los valores reales. Esto quiere decir, el error no es más que el total absoluto del error más no es el total del error al cuadrado. Por lo tanto, el error absoluto relativo toma el total y absoluto error que se normaliza dividiendo por el total de error absoluto de la predicción simple. Matemáticamente, el error relativo absoluto E_i de un individuo i es evaluado por la siguiente ecuación:

$$E_i = \frac{\sum_{j=1}^n |P_{(ij)} - T_j|}{\sum_{j=1}^n |T_j - \bar{T}|}$$

Figura 2.11: Relative Absolute Error.

Dónde:

$P_{(ij)}$: Es el valor que se ha esperado para el valor i del caso j (fuera de los n casos simples); T_j es el valor objetivo para el caso j ; y \bar{T} esta dado por la fórmula:

$$\bar{T} = \frac{1}{n} \sum_{j=1}^n T_j$$

Figura 2.12: Fórmula \bar{T} .

De mejor manera, el numerador es igual a 0 y a $E_i = 0$. Por tanto el índice E_i va desde 0 hasta infinito, donde 0 corresponde al ideal (Universidad Autónoma del Estado de México , 2010).

e. Error Cuadrático Relativo(Root Relative Squared Error):

UCLM resalta que error cuadrático relativo: “es el total de error cuadrático hecho relativo a lo que el error habría sido si la predicción fuese el promedio del valor absoluto.

Fórmula simple no es más que la media de los valores reales. De este modo, la relativa de error al cuadrado toma el total de errores al cuadrado y se normaliza dividiendo por el total de errores simples al cuadrado. Al tomar la raíz cuadrada del valor relativo, el error se reduce a las mismas dimensiones que la cantidad prevista.

Matemáticamente, la raíz cuadrada de error relativo E_i de un individuo i es evaluado por la siguiente ecuación:

$$E_i = \sqrt{\frac{\sum_{j=1}^n (P_{(ij)} - T_j)^2}{\sum_{j=1}^n (T_j - \bar{T})^2}}$$

Figura 2.13: Root Relative Squared Error.

Dónde:

$P_{(ij)}$: Es el valor que se ha esperado para el valor i del caso j (fuera de los n casos simples); T_j es el valor objetivo para el caso j ; y \bar{T} está dado por la fórmula de la Figura 2.12” (uclm, 2012).

f. Número total de instancias(Total Number of Instances):

María García Jiménez y Aránzazu Álvarez Sierra prevalencen que el número total de instancias “es el número de ejemplos que se han utilizado en la predicción” (García Jiménez & Álvarez Sierra , 2013).

2.2.9. Metodologías de la Minería de Datos

Juan Miguel Moine resalta que: “una metodología consiste en un conjunto de actividades organizadas que tienen por objetivo la realización de un trabajo. Para cada actividad se define las entradas y salidas, la forma en la que debe llevarse a cabo. Las metodologías permiten llevar a cabo el proceso de minería de datos en forma sistemática y no trivial. Ayudan a las organizaciones a entender el proceso de descubrimiento de conocimiento y proveen una guía para la planificación y ejecución de los proyectos.

Algunos modelos conocidos como metodologías son en realidad un modelo de proceso: un conjunto de actividades y tareas organizadas para llevar a cabo un trabajo. La diferencia fundamental entre metodología y modelo de proceso radica en que el modelo de proceso establece qué hacer, y la metodología especifica cómo hacerlo. Una metodología no solo define las fases de un proceso sino también las tareas que deberían realizarse y cómo llevar a cabo las mismas.

A continuación se mencionan las metodologías más importantes de la minería de datos:

- a. KDD(Knowledge Discovery in Databases).**
- b. Metodología SEMA.**

c. Metodología Crisp-DM.

La Metodología CRISP-DM profundiza en mayor detalle sobre las tareas y actividades a ejecutar en cada etapa del proceso de minería de datos, mientras que otros proveen sólo una guía general del trabajo a realizar en cada fase (como el proceso KDD o SEMMA)” (Moine J. M., Metodologías para el descubrimiento de Conocimiento en Bases de Datos, 2013).

2.2.10. Metodología CRISP-DM

IBM resalta que: “CRISP-DM son las siglas de Cross-Industry Standard Process for Data Mining, es un método probado para orientar los trabajos de minería de datos” (IBM, 2012).

Juan Miguel Moine, Ana Silvia Haedo y Silvia Gordillo resaltan que: “CRISP-DM fue creada por el grupo de empresas: SPSS, NCR y Daimler Chrysler en el año 2000, es actualmente la guía de referencia más utilizada en el desarrollo de proyectos de Data Mining. Estructura el proceso en seis fases: Comprensión del negocio, Comprensión de los datos, Procesamiento de los datos, Modelado, Evaluación y Despliegue. La sucesión de fases, no es necesariamente rígida. Cada fase es descompuesta en varias tareas generales de segundo nivel. Las tareas generales se proyectan a tareas específicas, pero en ningún momento se propone como realizarlas. Es decir, CRISP-DM establece un conjunto de tareas y actividades para cada fase del proyecto pero no especifica cómo llevarlas a cabo” (Moine , Haedo , & Gordillo, 2012).

IBM resalta que: “es un método probado para orientar trabajos de minería de datos. Y Como metodología, incluye descripciones de las fases normales de un proyecto, las tareas necesarias en cada fase y una explicación de las relaciones entre las tareas. Y como modelo de proceso, CRISP-DM ofrece un resumen del ciclo vital de minería de datos.

CRISP-DM organiza el desarrollo de un proyecto de Minería de Datos en una serie de fases o etapas, con tareas generales y específicas que permitan cumplir con los objetivos del trabajo de Investigación. Estas fases funcionan de manera cíclica e iterativa, pudiendo regresar desde alguna fase a otra anterior. En la Figura 2.14 se presenta las fases propuestas por esta metodología” (IBM, 2012):

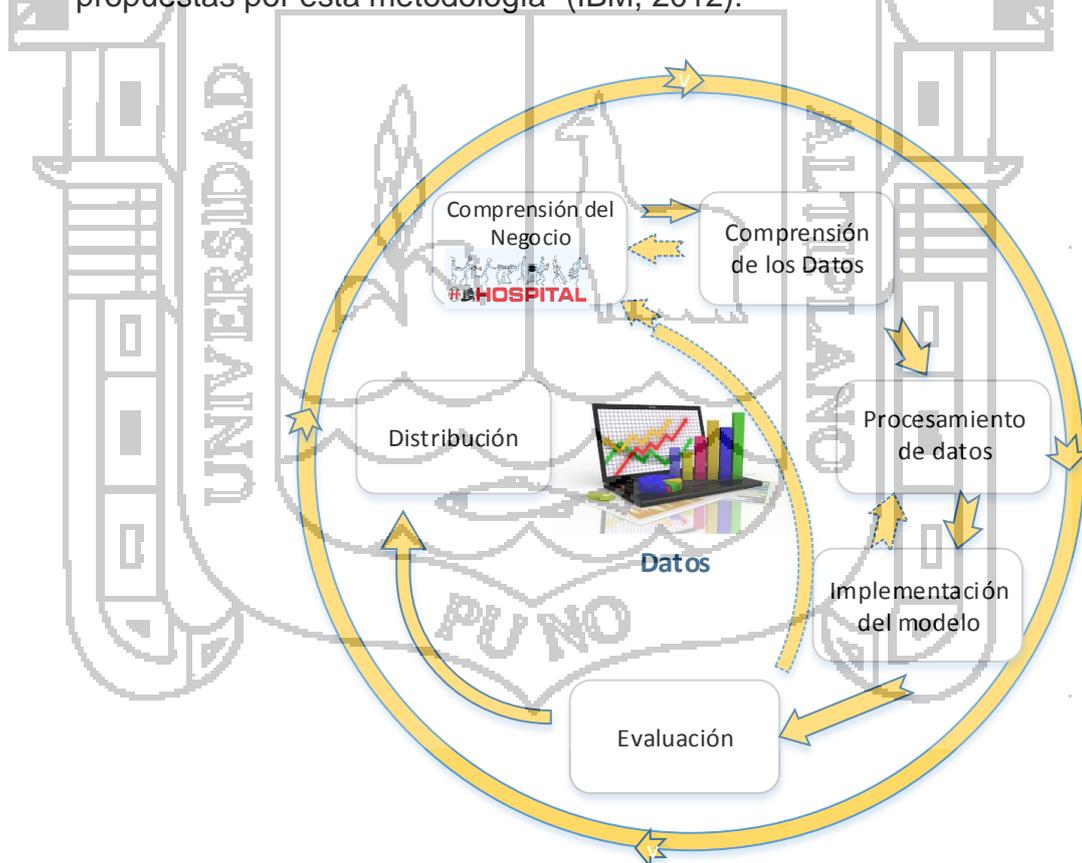


Figura 2.14: Fases de la Metodología CRISP-DM.

2.2.11. Fases de la Metodología CRISP-DM

Juan Manuel Moine a continuación define brevemente las fases y las subfases de cada fase (Moine J. M., 2013):

a. Comprensión del Negocio:

Marcelo Barrios resalta que: “en esta fase se determinan los objetivos y requerimientos del proyecto desde una perspectiva del negocio, definiendo el problema de minería y el plan de trabajo” (Barrios, 2010).

▪ Determinación de Objetivos de Negocio:

CEI rescata que: “el primer objetivo para un contexto es entender, desde una perspectiva de negocio, lo que el cliente realmente quiere lograr. A menudo el cliente tiene muchos objetivos que compiten y restricciones que deben ser correctamente equilibrados. En el objetivo se debe mostrar factores importantes.

En esta tarea se debe entender y establecer cuáles son los objetivos que el cliente pretende alcanzar, desde una perspectiva del negocio. Lo que se debe obtener es:

- Background, con la información que se conoce sobre la situación actual de la organización, incluyendo una descripción general del problema y la solución actual del mismo (si es que existe).
- Objetivos del negocio, identificando los objetivos principales del cliente.
- Criterios de éxito, describiendo los resultados esperados desde una perspectiva de negocio” (CEI, 2010).

▪ **Evaluación de la Situación:**

Xavier Roegiers y Alexia Peyser prevalecen que: “Esta tarea implica la investigación más detallada sobre todos los recursos, restricciones, presunciones, y otros factores que deberían ser considerados en la determinación del objetivo de análisis de datos y el plan de proyecto. En la tarea anterior, su objetivo es para ponerse rápidamente al porqué de la situación.

En esta tarea se debe profundizar en la evaluación de la situación actual del negocio. Analizar con mayor profundidad las restricciones y factores que se deben tener en cuenta para el proyecto. Lo que se debe obtener es lo siguiente:

- Inventario de recursos, donde deberán incluirse los recursos disponibles (como los recursos humanos, fuentes de datos, hardware y software).
- Lista de requerimientos del proyecto, supuestos y restricciones que se han detectado.
- Riesgos y planes de contingencia. Consiste en la identificación de los potenciales riesgos para el proyecto y la planificación de las acciones reactivas que se llevarán a cabo (planes de contingencia).
- Glosario con terminología relevante para el proyecto. En el mismo deberá incluirse un glosario de terminología del negocio y otro de minería de datos.
- Análisis costo-beneficio del proyecto” (Roegiers & Peyser, 2010).

- **Determinación de los Objetivos de la Minería de Datos:**

Paola Britos resalta que: “un objetivo de negocio declara objetivos en la terminología de negocio. Un objetivo de minería de datos declara objetivos de proyecto en términos técnicos. Por ejemplo, el objetivo de negocio podría ser “Aumentar catálogos de ventas a clientes existentes.”

Un objetivo de minería de datos podrían ser “Predecir cuantas baratijas un cliente comprará, obteniendo datos de sus compras de tres años pasados, información demográfica (edad, sueldo, ciudad y otros.), y el precio del artículo.” Lo que se obtiene con esta tarea es lo siguiente:

- Objetivos de minería de datos, describiendo los resultados previstos del proyecto que permiten el logro de los objetivos de negocio.
- Definir un criterio de éxito para el proyecto de minería. Especificar las condiciones bajo las cuales se aceptaran los resultados obtenidos” (Britos , 2010).

- **Crear un Plan para el Proyecto de Minería de Datos:**

Paola Britos rescata que: “Describir el plan intencionado para alcanzar los objetivos de minería de datos y así alcanzar los objetivos de negocio.

El plan debe especificar los pasos para ser realizados durante el resto del proyecto, incluyendo la selección inicial de herramientas y técnicas. El crear una planificación para el proyecto de minería debe ser consistente con los objetivos planteados. En esta tarea se obtiene lo siguiente:

- Plan de proyecto: listar las tareas que deben ser ejecutadas, duraciones y recursos necesarios, así como sus entradas y sus salidas.

El plan de proyecto es un documento dinámico que debe ser ajustado y revisado en cada fase.

- Evaluación inicial de técnicas y herramientas de minería que podrían ser utilizadas en el proyecto” (Britos , 2010).

b. Comprensión de los Datos

Pilar Folgueiras Bertomeu resalta que: “esta fase consiste en la recolección de datos que se utilizarán en el proyecto y la familiarización con los mismos. En esta etapa es posible el surgimiento de las primeras hipótesis acerca de la información que podría estar oculta y a continuación se detallan las tareas que se deben cumplir.” (Folgueiras Bertomeu, 2010).

- **Recolección de Datos Iniciales:**

CEI resalta que: “Esta recolección inicial incluye carga de datos, si es necesario para la comprensión de los datos. En esta tarea se recolecta todos los datos necesarios especificados en la lista de recursos del proyecto lo que se debe obtener con esta tarea es:

Reporte de recolección inicial de datos, donde se detalla la forma en la que han sido obtenidos los conjuntos de datos y los problemas que han surgido (CEI, 2010).

- **Describir los Datos:**

Barrios, Marcelo rescata de esta tarea lo siguiente: “se describe en líneas generales los datos recolectados y se obtiene lo siguiente:

Descripción de los datos, incluyendo el formato de los mismos y su tamaño (como cantidad de registros y variables)” (Barrios, 2010).

- **Explorar los Datos:**

Barrios, Marcelo resalta que: “Se realiza una exploración de los datos, observando la distribución y comportamiento de las variables con mayor relevancia.

En esta fase es conveniente el uso de técnicas simples de análisis estadístico en el que se obtiene lo siguiente:

Reporte inicial de exploración de datos, donde se expongan los resultados del análisis y la hipótesis iniciales con su impacto en el proyecto (Barrios, 2010).

- **Verificar la Calidad de los Datos:**

Pilar Folgueiras Bertomeu resalta que: “Examinar la calidad de los datos, incluyendo un análisis de su completitud, de potenciales errores en los mismos y de los datos ausentes. Se examina la calidad de los datos, dirigiendo preguntas como: ¿Los datos están completos? (¿Esto cubre todo los casos requeridos)? ¿Son correctos, o estos contienen errores y, si hay errores, que tan comunes son estos? ¿Hay valores omitidos en los datos? Si es así, ¿cómo se representan estos, donde ocurre esto, y que tan comunes son estos?. Lo que se obtiene en esta tarea:

Reporte de calidad de los datos, donde se documenta el análisis de calidad efectuado y propongan potenciales soluciones a los problemas encontrados” (Folgueiras Bertomeu, 2010).

c. Procesamiento de Datos

Rigeiro rescata que procesamiento de los datos: “comprende aquellas actividades de tratamiento de los datos para construir la vista minable o conjunto de datos final sobre el cual se aplicarán las técnicas de minería.

La fase de preparación de datos cubre todas las actividades necesarias para construir el conjunto de datos final (los datos que serán provistos en las herramientas de modelado) de los datos en brutos iniciales. Las tareas de preparación de datos probablemente van a ser realizadas muchas veces y no en cualquier orden prescripto. Las tareas incluyen la selección de tablas, registros, y atributos, así como la transformación y la limpieza de datos para las herramientas que modelan y se definen a continuación” (Rigeiro, 2012).

- **Selección de Datos:**

UNED rescata que: “En esta etapa se debe seleccionar con qué atributos (columnas) y con qué observaciones (filas o registros) se trabajará. La selección debe estar justificada lo que se obtiene:

Documento donde se justifiquen las causas por las cuales se incluyeron o excluyeron los datos” (UNED, 2012).

- **Limpieza de Datos:**

UNED define que: “Es una etapa que tiene por objetivo mejorar la calidad de los datos. En ella se deberán tomar decisiones sobre los problemas de calidad encontrados en los mismos, como datos ausentes o datos anómalos. En esta tarea la salida sería la siguiente:

Reporte de limpieza de datos, donde se incluyan las decisiones tomadas sobre los problemas de calidad de los datos (reportados en la fase Verificar la calidad de los datos”)” (UNED, 2012).

▪ **Construir Datos:**

José Luis Llavona Arregui rescata que: “En esta fase se lleva a cabo la construcción de nuevos datos, derivados de los disponibles, que son importantes para el análisis. Estos nuevos datos pueden ser, por ejemplo, atributos calculados o atributos transformados. Lo que se debe obtener es lo siguiente:

- Atributos derivados. Estos atributos se calculan a partir de otros atributos del mismo registro. Por ejemplo: $\text{edad_paciente} = \text{fecha_atencion} - \text{fecha_nacimiento}$.

▪ Registros creados. Estos nuevos registros se crean cuando son necesarios en la fase posterior del modelado (Llavona Arregui, 2010).

▪ **Integrar Datos:**

José Luis Llavona Arregui rescata que: “Esta tarea consiste en la integración de datos provenientes de diferentes tablas o registros. La salida de esta tarea es la siguiente:

Datos combinados. Resultan de integrar la información de dos o más tablas que tienen diferente información de las mismas observaciones. Por ejemplo, la integración de los datos personales y los datos de las atenciones efectuadas a un paciente en el departamento de cirugía. En esta fase se incluye el cálculo de agregaciones, donde se calculan nuevos

datos resumiendo información de diferentes tablas y registros. Siguiendo con el ejemplo del departamento de cirugía, podríamos integrar en un solo registro los datos personales del paciente, el total de atenciones efectuadas, y el promedio anual de operaciones realizadas” (Llavona Arregui, 2010).

- **Formatear Datos:**

UNED rescata que: “Esta etapa se refiere al cambio que debe realizarse e en formato de los datos (pero no en su significado) por los requisitos de las técnicas de modelado elegidas. Por ejemplo, el formato de las fechas o el ordenamiento del set de datos la salida es lo siguiente: Conjunto de fechas reformateados” (UNED, 2012).

- d. Implementación del Modelo**

William Meredith resalta que: “en esta etapa se aplican las diversas técnicas y algoritmos de minería sobre el conjunto de datos para obtener la información oculta y los patrones implícitos en ellos se cumplen las tareas siguientes en que también se definen cada un de ellos a continuación:” (Meredith, 2015).

- **Selección de la Técnica de Modelado:**

Guillermo Gilberto Molero castillo Resalta que: “esta tarea consiste en seleccionar qué técnica de minería de datos será utilizada. Por ejemplo, en un la investigación se utilizó la técnica de regresión lineal. Si se ha optado por el uso de múltiples técnicas, se debe repetir esta tarea para cada una. Las salidas de esta tarea son:

- La Técnica de modelado. Documentar la técnica de modelado con la que se trabajará.
- Supuestos del modelo. Algunas técnicas asumen supuestos sobre el conjunto de datos, como por ejemplo distribución normal de una variable. Documentar todos los supuestos realizados” (Molero Castillo, 2008).

- **Diseñar las Pruebas del Modelo:**

Regeiro rescata que: “Una vez contruidos los modelos, necesitaremos un mecanismo para determinar su calidad y validez. Por ejemplo, en problemas de agrupamiento se puede utilizar el coeficiente de silueta para evaluar la robustez de los grupos encontrados y en problemas de clasificación la tasa de error para estimar la capacidad del clasificador. En esta fase se dividirá el conjunto de datos en un grupo para entrenar el modelo (training) y otro para probarlo (test). La salida de esta tarea es:

El diseño de los test. Determinar y documentar de qué forma se entrenaran y evaluaran los modelos generados. Incluir las decisiones tomadas sobre los datos que se utilizaran para entrenamiento y prueba” (Rigeiro, 2012).

- **Construir el Modelo:**

Guillermo Gilberto Molero Castillo rescata que: “en esta subfase se debe aplicar la técnica seleccionada sobre el conjunto de datos para generar uno o más modelos, también el modelo será evaluado con distintos valores de parámetros. Las salidas de la siguiente tarea son:

- Parámetros seleccionados. Listar los parámetros que se han proporcionado al modelo, justificando la elección de los mismos.
- Modelos producidos por las herramientas de minería.
- Descripción de los modelos” (Molero Castillo, 2008).

▪ **Evaluación del Modelo:**

Pablo Romeau Guallart rescata que: “En esta fase, el equipo de proyecto interpreta y evalúa el modelo en función de su conocimiento del dominio, los criterios de éxito definidos para el proyecto (tarea determinar los objetivos de minería de datos) y las pruebas diseñadas para el modelo (tarea diseñar las pruebas del modelo). Los modelos pueden ser valorados y renqueados” (Romeau Guallart, 2010).

e. Evaluación

William Meredith resalta que: “es la Fase en la que se analizan los patrones obtenidos en función de los objetivos organizacionales. En esta etapa se debería determinar si se ha omitido algún objetivo importante del negocio y si el nuevo conocimiento será implementado, es decir, si se pasará a la próxima etapa y a continuación se definen las siguientes tareas que se deben cumplir.” (Meredith, 2015).

▪ **Evaluación de los Resultados**

Daycy Elizabeth Peralta Cochancela resalta que: “En esta etapa se evalúa el modelo en función de los objetivos del negocio, determinando su validez de acuerdo a los intereses organizacionales. Además del modelo, puede haber surgido como parte del proceso nueva información relevante y futuras líneas de investigación.

- Evaluación de los resultados de la minería de datos con respecto a los criterios de éxito y objetivos de negocio.
- Modelos evaluados y aprobados” (Peralta Cochancela, 2009).

▪ **Revisión del Proceso:**

Federico Peralta rescata que: “en esta etapa realizar una revisión completa del proceso efectuado en búsqueda de posibles errores u omisiones. La salida de esta tarea es:

La revisión del proceso, documentando un resumen del mismo. Incluir las actividades omitidas o bien aquellas que deberían ser repetidas (Peralta, 2014).

▪ **Determinación de los Próximos Pasos:**

UNED rescata que: “En función de la evaluación de resultados y la revisión del proceso, se debe decidir cómo continua el proyecto: si se pasa a la próxima fase (implementación) o bien si se retorna a una fase anterior. Las salidas de esta tarea son:

- Lista de posibles acciones.
- Descripción de la decisión tomada” (UNED, 2012).

f. Distribución

José Hernández Orallo y César Ferri Ramírez resaltan que: “Consiste en la comunicación e implementación del nuevo conocimiento, el cual debe ser representado de forma entendible las fases que se deben cumplir son:

Planificar la implementación, planificar el monitoreo y el mantenimiento, crear un reporte final, revisar el proyecto a continuación se define cada una de las tareas:” (Hernández Orallo & Ferri Ramírez, 2013).

- **Planificar la implementación**

Juan Miguel Moine rescata que: “En esta etapa se genera el plan de implementación de los resultados obtenidos mediante la minería de datos. La salida de esta tarea es:

Plan de implementación, incluyendo las etapas y como llevarlas a cabo (Moine J. M., Metodologías para el descubrimiento de conocimiento en bases de datos: un estudio comparativo, 2013).

- **Planificar el Monitoreo y el Mantenimiento**

Juan Miguel Moine rescata que: “El monitoreo y mantenimiento es de gran importancia si los resultados de la minería formarán parte del trabajo diario del negocio y su entorno. Su salida es:

El plan de mantenimiento y monitoreo” (Moine J. M., Metodologías para el descubrimiento de conocimiento en bases de datos: un estudio comparativo, 2013).

- **Crear un reporte final**

Pilar Folgueiras Bertomeu resalta que: “en esta etapa generar un reporte final, que podría resumir el desarrollo del proyecto o bien mostrar un análisis comprensivo de los resultados obtenidos en el proceso de minería.

Las salidas de esta tarea se muestran a continuación:

- Reporte final del proyecto.
- Presentación final al cliente, incluyendo resultados y conclusiones”
(Folgueiras Bertomeu, 2010).

▪ **Revisión del Proyecto**

Meredith, William rescata que: “esta subfase consiste en identificar y analizar los puntos que fueron bien realizados, los que fueron mal realizados, y los que podrían mejorarse.

Documentación de la experiencia adquirida durante el desarrollo del proyecto” (Meredith, 2015).

2.2.12. Software para Minería de Datos:

Isabel Ramos Román y José Javier Dolado Cosín recatan que: “Algunos de los algoritmos que más se usan fueron creados hace 30 años, lo cual hace que hoy existen productos de software para minería de datos que generan resultados de alta confiabilidad” (Ramos Román & Dolado Cosín, 2007).

Francisco José García González rescata que: “los Software de la Minería de Datos se clasifican en y se detallan a continuación:” (García González, 2013):

a. Comerciales

Los Softwares de Minería de Datos comerciales más utilizados son los siguientes:

- Intelligent Miner /DB2 Data Warehouse Edition (IBM).

- Clementine.
- Enterprice Miner (SAS).
- DataEngine.
- Oracle Datamining.

b. De Código Libre

Existen varias aplicaciones de uso libre, entre las más usadas se encuentra:

- Weka.
- RapidMiner.
- Knime.

2.2.13. Weka

Luis Córdoba Falla Rescata que: “Weka se trata de un acrónimo derivado de Waikato Environment for Knowledge Analysis – Entorno para Análisis del Conocimiento de la Universidad de Waikato. Weka es un software programado en Java que está orientado a la extracción de conocimientos desde bases de datos con grandes cantidades de información” (Córdoba Fallas, 2011).

Agustín José Calleja Gómez resalta que: “weka es una potente herramienta de minería de datos que se encuentra libremente disponible bajo la licencia pública general de GNU, lo que significa que este programa es de libre distribución y difusión ya que funciona en cualquier plataforma sobre la que haya una máquina virtual Java disponible. Weka tiene cuatro opciones posibles de trabajo, como se muestran en la Figura

2.15, y que a continuación se describen brevemente:" (Calleja Gómez, 2010):

a. Simple CLI

La interfaz "Command-Line Interfaz" es simplemente una ventana de comandos java para ejecutar las clases de WEKA.

b. Explorer

Es la opción que permite llevar a cabo la ejecución de los algoritmos de análisis implementados sobre los ficheros de entrada, una ejecución independiente por cada prueba. Esta es la opción sobre la que se hizo centra la totalidad de esta guía.

c. Experimenter

Esta opción permite definir experimentos más complejos, con objeto de ejecutar uno o varios algoritmos sobre uno o varios conjuntos de datos de entrada, y comparar estadísticamente los resultados.

d. KnowledgeFlow

Esta Opción permite llevar a cabo las mismas acciones del "Explorer", con una configuración totalmente gráfica, para seleccionar componentes y conectarlos en un proyecto de minería de datos, desde que se cargan los datos, se aplican algoritmos de tratamiento y análisis, hasta el tipo de evaluación deseada.

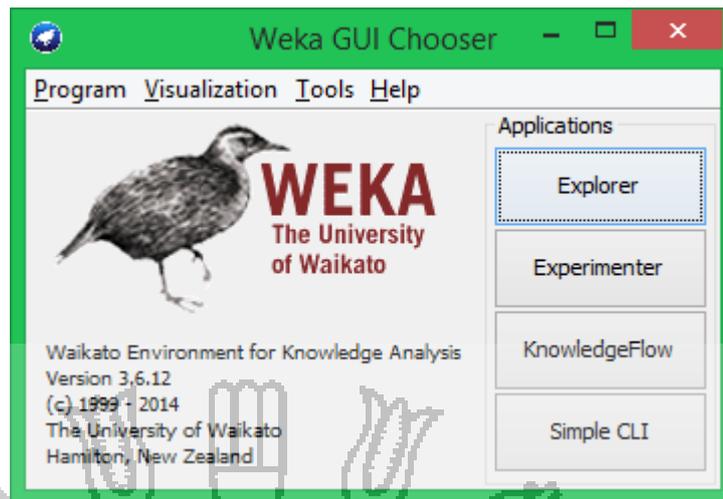


Figura 2.15: Opciones de Trabajo de Weka.

Ricardo Aler rescata que: “en la opción Explorer una vez seleccionada, se crea una ventana con 6 pestañas en la parte superior con diferentes tipos de operaciones, en etapas independientes, que se pueden realizar sobre los datos:

- Preprocess: Selección de la fuente de datos y preparación (filtrado).
- Clasify: Facilidades para aplicar esquemas de clasificación, entrenar modelos y evaluar su precisión.
- Cluster: Algoritmos de agrupamiento.
- Associate: Algoritmos de búsqueda de reglas de asociación.
- Select Attributes: Búsqueda supervisada de subconjuntos de atributos representativos.
- Visualize: Herramienta interactiva de presentación gráfica en 2D” (Aler, 2009).

El preprocesado de los datos es la primera parte por la que se debe pasar antes iniciar con cualquier otra operación, ya que se precisan

datos para poder llevar a cabo cualquier análisis. La disposición de la parte de preprocesado del Explorer, Preprocess, se indica en la Figura 2.16.

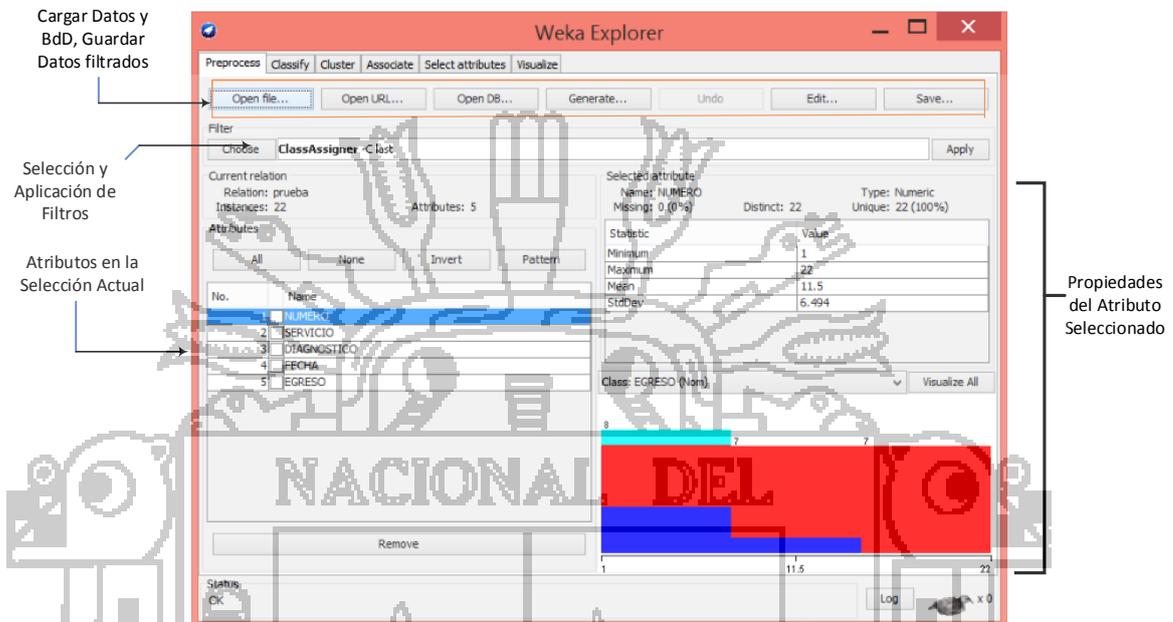


Figura 2.16: Preprocesado de los Datos.

2.2.14. Toma de Decisiones

Roberto García Reyes rescata que: “Toma de decisión es la elección entre alternativas múltiples, usualmente hechas con justo rango de raciocinio. Asimismo la toma de decisiones forma parte de una estrategia de resolución de problemas y se refiere al proceso mediante el cual un grupo de individuos trata de acortar la distancia que existe entre las condiciones de operación actuales del sistema y las condiciones mejoradas supuestas, que se desean alcanzar en un futuro.

Para entender los problemas que se enfrentan al tomar una decisión se debe tomar en cuenta el número de factores involucrados o indicadores.

La toma de una decisión correcta implicará la selección de la alternativa que cumpla el mejor criterio de operación entre todas las alternativas existentes” (Reyes, 2012). Ver la Figura 2.17 .

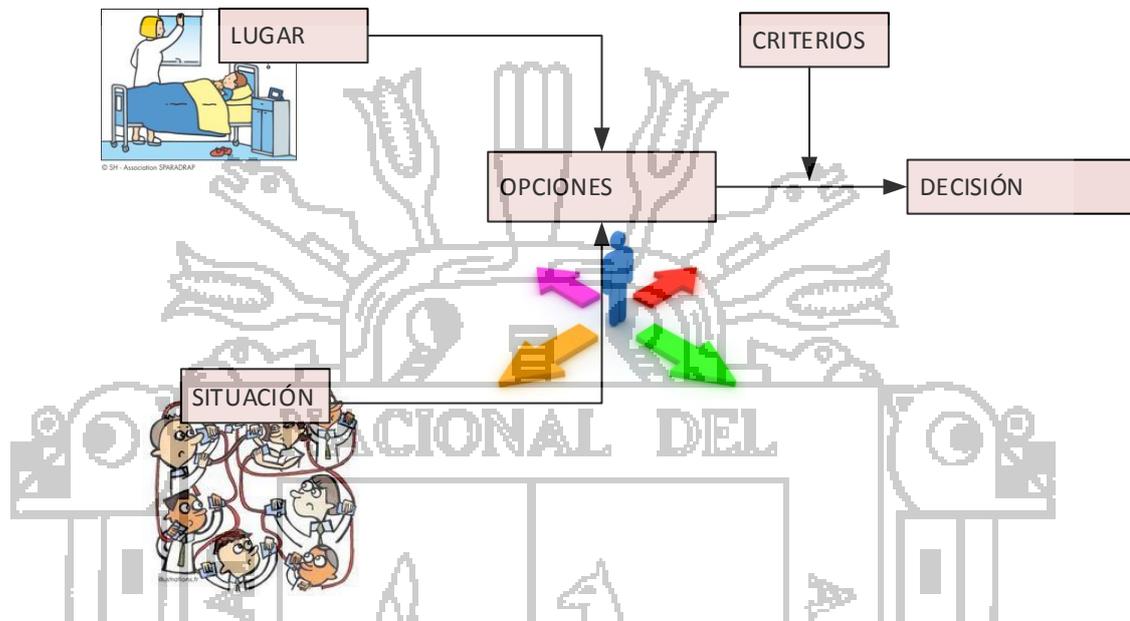


Figura 2.17: Estructura de la Toma de Decisiones.

2.2.15. Toma de Decisiones en el Departamento de Cirugía

Según el Hospital Regional Manuel Núñez Butrón para: “ la mejora en la toma de decisiones en el departamento de cirugía los indicadores hospitalarios son los más importantes estos deben ser conceptualmente claros, directos, inequívocos y comprensibles, que signifiquen lo mismo tanto a quienes hacen uso de ellos con fines de toma de decisiones, como a quienes simplemente los observan como referencia por ello se definen los indicadores siguientes para el Departamento de Cirugía (Hospital Manuel Núñez Butrón, 2013):

a. Número de Intervenciones Quirúrgicas

Es el resultado de sumar las operaciones programadas o electivas y de emergencia, así como las llamadas y todas aquellas que se lleven a cabo en el centro quirúrgico.

Número de Intervenciones Quirúrgicas por Servicio o Especialidad: Es el producto de sumar las operaciones por separado de cada uno de los servicios o especialidades quirúrgicas.

$$NIQ = \sum OP$$

Figura 2.18: Número de Intervenciones Quirúrgicas

Dónde:

NIQ: Número de Intervenciones Quirúrgicas

OP: Operaciones programadas

b. Número de Pacientes Operados:

Paciente operado: Es aquel en quien se ha practicado una o más intervenciones quirúrgicas en el centro quirúrgico del hospital, es el producto de sumar todas las personas operadas en los servicios, aunque puede desagregarse por cada uno de ellos.

$$NPO = \sum PO$$

Figura 2.19: Número de Pacientes Operados

Dónde:

NPO: Número de Pacientes Operados

PO: Pacientes operados

c. Aprovechamiento de Salas de Operaciones:

Sirve para medir el grado real de utilización de la capacidad instalada del centro. Se utilizan los siguientes parámetros para medir:

Horas Quirúrgicas Efectivas o de Uso de Sala Quirúrgica: Es la suma del tiempo real, de utilización de sala de operaciones por cada intervención quirúrgica, el cual se contabiliza desde el ingreso del paciente hasta su salida de la sala de emergencia.

Horas Quirúrgicas Disponibles: Es el tiempo que puede utilizarse para las intervenciones quirúrgicas, valor que está condicionado por la cantidad de salas de operaciones del hospital.

$$ASO = \frac{NHQEP}{NHQDMP}$$

Figura 2.20: Aprovechamiento de Salas de Operación

Dónde:

ASO: Aprovechamiento de salas de operaciones.

NHQEP: N° de horas quirúrgicas efectivas en un periodo.

NHQDMP: N° de horas quirúrgicas disponibles en el mismo periodo.

d. Grado de Cumplimiento de las Operaciones Programadas:

Este indicador establece el porcentaje de intervenciones quirúrgicas realizadas respecto a las programadas, es decir, mide el grado de cumplimiento de las programaciones de operaciones electivas. Por lo tanto, no se deben incluir las intervenciones de emergencia.

Número de Intervenciones Quirúrgicas Programas Ejecutadas: Es el resultado de sumar las programadas que han sido llevadas a cabo,

incluyendo las llamadas cirugías de día y todas aquellas que utilicen el centro quirúrgico, excluyendo las de emergencia.

Número de Intervenciones Quirúrgicas Programadas: Es el resultado de sumar las que fueron programadas en los turnos quirúrgicos regulares y extraordinarios.

$$GCOP = \frac{NIQPEjP}{NIQPMP} * 100$$

Figura 2.21: Grado de Cumplimiento de la Operaciones Programadas.

Dónde:

GCOP: N° de intervenciones quirúrgicas programadas ejecutadas en un periodo.

NIQPEP: N° de intervenciones quirúrgicas programadas ejecutadas en un periodo.

e. Rendimiento del Quirófano:

Es la relación del número de intervenciones quirúrgicas realizadas y el número de sala de operaciones del establecimiento de salud en un determinado periodo de tiempo, mide el número promedio de intervenciones quirúrgicas realizadas por cada sala de operaciones.

$$RQ = \frac{NIQEj}{NSO} * 100$$

Figura 2.22: Rendimiento del Quirófano.

Dónde:

RQ: Rendimiento de quirófano

NIQEj: N° de intervenciones quirurgicas ejecutadas

NISO: N° de salas de operaciones

f. Porcentaje de Intervenciones Quirúrgicas Suspendidas:

Es el porcentaje de cirugías suspendidas con respecto a las cirugías programadas, mide el grado de cumplimiento de las actividades quirúrgicas programadas. La suspensión de estas actividades puede estar asociada a la falta de material y/o instrumental médico.

Número de Intervenciones Quirúrgicas Pendientes: Mide la cantidad de pacientes que se encuentran en lista de espera para ser intervenidos quirúrgicamente en cada servicio o especialidad. Permite tener una idea de la presión de la demanda que existe sobre el centro quirúrgico.

$$PIQS = \frac{NIQSP}{NIQPMP} * 100$$

Figura 2.23: Porcentaje de Intervenciones Quirúrgicas Suspendidas.

Dónde:

PIQS: Porcentaje de intrvenciones quirúrgicas suspendidas.

NIQSP: N° de intervenciones quirúrgicas suspendidas en un periodo.

NIQPMP: N° de intervenciones quirúrgicas programadas en el mismo periodo.

g. Promedio de Días Preoperatorios:

Sirve para medir la utilización eficiente de las camas hospitalarias de los servicios quirúrgicos, permite conocer los días cama disponibles mal utilizados por pacientes en espera de una intervención quirúrgica, el periodo preoperatorio de utilización de camas debe ser el menor posible y debe estar en relación a la capacidad instalada del centro quirúrgico para evitar que las camas sean ocupadas innecesariamente.

Días Preoperatorios: Son aquellos durante los cuales permanece hospitalizado un paciente, previos a la intervención quirúrgica, el número de días preoperatorios es la suma de los días de espera de todos los pacientes operados en un periodo.

$$PDP = \frac{NDPrePO}{NPOMP}$$

Figura 2.24: Promedio de Días Preoperatorios.

Dónde:

PDP: Promedio de días preoperatorios.

NDPrePO: N° de días preoperatorios de pacientes operados.

NPOMP: N° de días preoperatorios de pacientes operados.

h. Porcentaje de Complicaciones Intraoperatorias:

Refleja la magnitud de las complicaciones que se presentan durante el acto quirúrgico, no se incluyen en el cálculo las complicaciones.

Complicación Intraoperatoria: Es un evento no deseado que se produce y se hace evidente durante la intervención quirúrgica el cual puede estar relacionado a las condiciones propias del paciente o a los profesionales de salud que participan dejando o no secuelas en el intervalo.

$$PCI = \frac{NCIP}{NOMP} * 100$$

Figura 2.25: Porcentaje de Complicaciones Intraoperatorias.

Dónde:

PCI: Porcentaje de complicaciones intraoperatorias.

NCIP: N° de complicaciones intraoperatorias en un periodo.

NOMP: N° de operaciones en el mismo periodo.

i. Porcentaje de Complicaciones Post Operatorias:

Sirve para establecer la proporción de complicaciones que se presentan luego de haber concluido al acto quirúrgico. No se incluyen en el cálculo las complicaciones originadas por la administración de la anestesia o por los procedimientos en la sala de recuperación.

Complicación Postoperatoria: Es un evento no deseado que puede originarse o no durante el acto quirúrgico, pero que se hace evidente luego de concluido este. Puede o no dejar secuelas en el paciente.

$$PCPO = \frac{NCIP}{NOMP} * 100$$

Figura 2.26: Porcentaje de Complicaciones Post Operatorias.

Dónde:

PCPO: Porcentaje de complicaciones Post Operatorias.

NCIP: Número de complicaciones intraoperatorias en un periodo.

NOMP: Número de operaciones en un mismo período.

j. Porcentaje de Complicaciones Anestésicas:

Mide la proporción de complicaciones asociadas a la administración de la anestesia durante el acto quirúrgico. No se incluyen las complicaciones relacionadas propias de la técnica quirúrgica.

Complicación Anestésica: Es un evento no deseado que se produce y se hace evidente durante o después de la intervención quirúrgica, el cual está directamente relacionado con la administración anestésica, ya sea

por la dosis o tiempo de exposición, las reacciones adversas a esa sustancia o la idiosincrasia del paciente puede o no dejar secuelas en el intervenido.

$$PCA = \frac{NCDA}{NPRAMP} * 100$$

Figura 2.27: Porcentaje de Complicaciones Anestésicas.

Dónde:

PCA: Porcentaje de complicaciones por anestésicas.

NCDA: N° de complicaciones debidas a anestesia.

NPRAMP: N° de pacientes que recibieron anestesia en el mismo periodo.

k. Porcentaje de Complicaciones en Recuperación:

Refleja la proporción de complicaciones que se presentan durante la estadía del paciente en la sala de recuperación.

Complicación en Recuperación: Es un evento no deseado que se produce y se hace evidente durante la permanencia del paciente operado el cual puede estar relacionado a las condiciones propias del paciente, procedimientos médicos o cuidados de enfermería durante su estancia en el servicio dejando o no secuelas en el intervenido. Se excluyen las causas asociadas al acto quirúrgico y a la administración anestesia.

La forma para calcular este indicador en la siguiente fórmula:

$$PCR = \frac{NCRDP}{PNPRMP} * 100$$

Figura 2.28: Porcentaje de Complicaciones en Recuperación.

Dónde:

PCR: Porcentaje de complicaciones en recuperación.

NCRDP: N° de complicaciones en recuperación durante un periodo.

NPRMP: N° de pacientes en recuperación en el mismo periodo.

I. Tasa de Mortalidad en el Centro Quirúrgico:

Este indicador permite evaluar la calidad de la atención en el Centro Quirúrgico, establece la proporción de fallecidos en este servicio por diversas causas, incluye a los pacientes fallecidos en sala de operaciones y de recuperación, así como a los fallecidos por administración anestésica.

Muerte o Defunción en el Centro Quirúrgico: Es la que se produce ya sea por el acto operatorio y sus complicaciones posteriores, o por la administración de anestesia y sus complicaciones posteriores.

$$TM = \frac{NDCQP}{NPPMP} * 100$$

Figura 2.29: Muerte o Defunción en el Centro Quirúrgico.

Dónde:

TM: Tasa de mortalidad.

NDCQP: N° de defunciones en el Centro Quirúrgico en un periodo.

NPPMP: N° de pacientes operados en el mismo periodo.

2.3. Glosario de Términos Básicos

2.3.1. Árboles de Decisión

Es una técnica útil para problemas en los que se presentan decisiones secuenciales. Aunque esta técnica es de mayor utilidad para situaciones en que el riesgo está presente también es empleada en condiciones de certeza (Liz, 2012).

2.3.2. Casos de Uso

Es una descripción de la secuencia de interacciones que se producen entre un actor y el sistema, cuando el actor usa el sistema para llevar a cabo una tarea específica. (Ferré Grau, Xavier; Sánchez Segura, María Isabel, 2011).

2.3.3. Dato

Es la unidad mínima de la información. Es un elemento aislado, recabado para un cierto fin, pero que no ha pasado por un proceso que lo interrelacione con otros de manera funcional para el fin previsto. Los datos recogen un conjunto de hechos (Molina López, 2006).

2.3.4. Diagramas de Secuencia

Resaltan la ordenación temporal de los mensajes que se intercambian (Booch, Grady; Rumbaugh, jim; Jacobson, Ivar, 2013).

2.3.5. Información

Es un conjunto de mecanismos que permiten a un individuo estructurarlos de una manera determinada. Los patrones, asociaciones o

relaciones entre los datos proporcionan información (Calderón Méndez, 2006).

2.3.6. Método

Es una serie de pasos sucesivos, conducen a una meta. El objetivo del profesionalista es llegar a tomar las decisiones y una teoría que permita generalizar y resolver de la misma forma problemas semejantes en el futuro (Eduardo, 2015).

2.3.7. Modelo Entidad Relación

Se basa en una percepción de un mundo real que consiste en un conjunto de objetos básicos llamados entidades y de relaciones entre estos objetos. Se desarrolló para facilitar el diseño de bases de datos permitiendo especificar un esquema empresarial. Este esquema representa la estructura lógica general de la base de datos (López Gaona, 2011)

2.3.8. Modelo de Minería de Datos

Es algo más que un algoritmo o un contenedor de metadatos: es un conjunto de datos, estadísticas y patrones que se pueden aplicar a los nuevos datos para generar predicciones y deducir relaciones (Microsoft, 2014).

2.3.9. Precisión

La norma ISO 3354 [ISO 1993] define la precisión como “el grado de concordancia entre ensayos independientes obtenidos bajo unas condiciones estipuladas”. Estas condiciones dependen de los factores

que se varíen entre cada uno de los ensayos. Medida como la proximidad existente entre las indicaciones o los valores medidos obtenidos en mediciones repetidas de un mismo objeto, o de objetos similares, bajo condiciones específicas (Maroto , Boqué , Riu, & Rius, 2011).

2.3.10. Predicción

Es un pronóstico razonable y verificable, sobre un hecho o acontecimiento nuevo o desconocido. Se caracteriza por anticipar lo que va a ocurrir; la inferencia, en cambio, explica o interpreta lo que ya ha ocurrido (Jankovic, 2011).

2.3.11. Técnicas

Son los procedimientos e instrumentos que se utilizan para acceder al conocimiento (Morone, 2013).

2.3.12. UML

Lenguaje Unificado de Modelado proporciona un vocabulario y unas reglas para permitir una comunicación. En este caso, este lenguaje se centra en la representación gráfica de un sistema. (Hernández Orallo E. , 2011).

2.3.13. Variables

Se refieren a aquellas propiedades, atributos, características, magnitudes, funcionalidades, etc. que podrán estar presentes (en algunos casos con distinta intensidad) o ausentes en cada uno de los casos que conforman el universo de estudio (Morone, 2013).

2.4. Hipótesis de la Investigación

El modelo predictivo basado en Minería de datos Mejora en la Toma de Decisiones del Departamento de Cirugía del Hospital Regional Manuel Núñez Butrón.

2.5. Operacionalización de Variables

- **Variable Independiente:** Modelo predictivo.
- **Variable dependiente:** Toma de decisiones.

Tabla 2.1: Operacionalización de Variables.

| Variables | Dimensión | Indicadores | Escala |
|-----------------------------------|---------------------------------|---|---|
| Independiente: Modelo Predictivo | Desarrollo de modelo predictivo | <ul style="list-style-type: none"> • Comprensión del negocio • Comprensión de los datos • Procesamiento de los datos • Implementación • Evaluación | 1: Realizado 0: No realizado |
| Independiente: Toma de decisiones | Evaluación | <ul style="list-style-type: none"> • Predicción • Datos Históricos | Malo 0-25% Regular 26-50% Bueno 51-75% Excelente 76-100% |

Fuente: Elaboración Propia.



**CAPÍTULO 3: DISEÑO METODOLÓGICO DE
INVESTIGACIÓN**

3.1. Tipo y Diseño de Investigación

La investigación de acuerdo a las características del problema, objetivo y la hipótesis es de tipo experimental, de diseño pre experimental, con pre-test y post-test, los diseños pre experimentales no presentan grupo control.

El grupo experimental estuvo conformado por los datos de los pacientes registrados en el periodo 2012-2013, en el departamento de Cirugía del Hospital Manuel Nuñez Butrón – Puno.

La representación gráfica es la siguiente:

$G1: 01 X 02$

Figura 3.1: Grupo Experimental.

Dónde:

$G1$: Grupo experimental.

X : Tratamiento con el modelo de minería de datos.

01 : Test antes del experimento.

02 : Test después del experimento.

Este diseño con grupo experimental permitió la comparación de resultados pre test y post test, con un alto grado de probabilidad, ante la situación con modelo de minería de datos y sin modelo (variable independiente), ha sido un factor determinante para la toma de decisiones en la predicción de datos del Departamento de Cirugía del Hospital Manuel Nuñez Butrón (Variable dependiente).

3.2. Población y Muestra de Investigación

Población: Todos los datos de pacientes atendidos en las diferentes especialidades en el Hospital Manuel Núñez Butrón Puno.

Muestra: Todos los datos de Pacientes atendidos en el periodo (2012-2013) en el Departamento de cirugía del Hospital Manuel Núñez Butrón Puno.

3.3. Ubicación y Descripción de la Población

Se investigó el Departamento de Cirugía, del Hospital Manuel Núñez Butrón Puno.

3.4. Material Experimental

▪ Encuestas:

Se aplicó encuestas a la Unidad de Estadística con el fin de obtener información requerida por el departamento de Cirugía como se muestra en el Anexo 1.

▪ Fichas de observación:

Se elaboró una ficha de observación para registrar los datos del experimento con modelo y sin modelo se muestra en el Anexo 2.

3.5. Técnicas e Instrumentos para Recolectar Información

Se utilizaron fichas de observación y observación directa.

3.6. Técnicas para el Procesamiento y Análisis de Datos

Se utilizó la técnica de regresión lineal.

3.7. Procedimiento del Experimento

Se aplicaron las fases de la metodología CRISP–DM:

- Comprensión del negocio.
- Comprensión de los datos.
- Preparación de datos.
- Modelado.
- Evaluación.

3.8. Plan de Tratamiento de los Datos

Los datos se procesaron de forma cualitativa y cuantitativa.

Para el caso del desarrollo cualitativo del modelo se documentaron los diagramas y casos de uso de los procesos de negocio y herramientas de software desarrollados.

Para los datos cuantitativos se utilizaron tablas y gráficos mediante un programa de hoja de cálculo.

3.9. Diseño Estadístico para la Prueba de Hipótesis o Prueba de Hipótesis

Para la prueba de hipótesis se utilizó las comparaciones entre los datos reales y los que el modelo devuelve con lo cual se halló el promedio según las siguientes formulas:

a. Predicción:

$$\text{Precisión} = \frac{N^{\circ} \text{ de predicciones correctas}}{N^{\circ} \text{ total de predicciones}} \times 100$$

Figura 3.2: Fórmula de Precisión

b. Datos históricos

$$\text{Datos históricos} = \frac{\sum(\text{indicadores con modelo} - \sum \text{ndicadores sin modelo})}{\sum \text{indicadores sin modelo}} \times 100$$

Figura 3.3: Fórmula de Datos Históricos

c. Mejora en la toma de decisiones

De las formulas a y b se obtiene el porcentaje de mejora para la toma de decisiones, como se muestra:

$$\text{Mejora en la toma de decisiones} = \frac{\text{Precisión} + \text{Datos Históricos}}{2}$$

Figura 3.4 Fórmula de Mejora en la Toma de Decisiones



4.1. Comprensión de la Situación Actual

En este punto se comprendió la situación actual realizando la recolección de información y se analizó dicha información sobre los objetivos y requisitos del departamento de cirugía desde una perspectiva institucional. Para lo cual se desarrollaron las siguientes tareas:

- Identificación de los objetivos del departamento de cirugía.
- Evaluación de la situación.
- Determinación de los objetivos de la minería de datos.
- Elaboración del plan de minería de datos.

4.1.1. Objetivos del Departamento de Cirugía

El departamento de Cirugía es la encargada de brindar atención a los pacientes con procedimientos quirúrgicos para su recuperación y rehabilitación; depende de la dirección del hospital, se organiza en dos servicios asistenciales: servicio de cirugía general y servicio de cirugía especializada, según organigrama Figura 4.1 (Hospital Regional Manuel Núñez Butrón, 2014):



Figura 4.1: Organigrama Departamento de Cirugía

a. Servicio de Cirugía General:

Es la Unidad Orgánica de brindar la atención por medios quirúrgicos en cirugía general, depende del departamento de cirugía y tiene los siguientes objetivos funcionales generales:

- Brindar atención integral, en tratamiento de pacientes por medios quirúrgicos en cirugía general.
- Realizar el diagnóstico, tratamiento y rehabilitación del paciente mediante acto quirúrgico.
- Realizar acciones recuperativas y de promoción y prevención en el campo de salud.
- Proponer, ejecutar y evaluar protocolos y procedimientos de atención de cirugía general orientados a proporcionar un servicio eficiente y eficaz.
- Prevenir riesgos y daños, proteger y recuperar la salud y rehabilitar las capacidades de los pacientes, en la atención médica de cirugía general, según las normas y protocolos vigentes.
- Asegurar el cumplimiento de las normas de bioseguridad

b. Servicio de Cirugía de Especializada

Es la Unidad Orgánica de brindar atención a pacientes en cirugía especializada, depende del departamento de cirugía y tiene los siguientes objetivos:

- Brindar atención integral, en tratamiento de pacientes por medios quirúrgicos en cirugía general, cirugía laparoscópica, neurocirugía, otorrinolaringología, traumatología y ortopedia, urología, cirugía

pediátrica, oftalmología en coordinación con las unidades orgánicas competentes y según los protocolos y normas vigentes.

- Realizar procedimientos quirúrgicos según las especialidades existentes.
- Ejecutar acciones de promoción, protección y recuperación de la salud en los casos con problemas quirúrgicos especializados.
- Proponer, ejecutar y evaluar protocolos y procedimientos de atención de cirugía general orientados a proporcionar un servicio eficiente y eficaz.
- Cumplir normas de bioseguridad.
- Estimular orientar y monitorear la investigación, en el campo de su competencia, así como apoyar la docencia en el marco de los gobiernos correspondientes.

Sin embargo dichos objetivos no son posibles de alcanzar sin una adecuada gestión de la información requerida y dada la gran cantidad de información que se genera diariamente se planteó el uso de minería de datos

Se determinó que el criterio de éxito es el control de los indicadores hospitalarios del departamento de cirugía, los cuales son base para una adecuada toma de decisiones.

4.1.2. Evaluación de la Situación Actual

Para desarrollar la minería de datos, se tuvo un conocimiento previo de 8 meses observando la situación actual en el departamento de cirugía.

Seguidamente se determinó que la cantidad de datos necesaria deberán ser extraídos de al menos 1500 informes operatorios, lo cual fue factible.

La aplicación de minería de datos tuvo los siguientes costos y beneficios:

Tabla 4.1: Costos y Beneficios.

| Costo | Beneficio |
|----------------------------|---|
| Digitación manual de datos | Base de datos digitalizada |
| Tiempo de ingreso de datos | Reportes inmediatos |
| Capacitación de personal | Personal capacitado en minería de datos |
| Recursos tecnológicos | Procesos automatizados |

Fuente: Elaboración Propia.

4.1.3. Objetivos de la Minería de Datos

Se lograron determinar los objetivos de la minería de datos en base al objetivo del negocio que es toma de decisiones el cual depende del control de los indicadores hospitalarios como se determinó en el punto 2.2.16. y se muestran en la tabla siguiente:

Tabla 4.2: Determinación de los Objetivos de Minería de Datos.

| Objetivo del negocio | Objetivo de Minería de Datos | Criterio de Éxito |
|---------------------------------|--|---------------------------------------|
| Mejora en la toma de decisiones | Control de los indicadores hospitalarios | Grado de precisión mayor igual al 75% |

Fuente: Elaboración Propia

4.1.4. Plan de Proyecto

El proyecto fue realizado en un periodo de 15 semanas, teniendo como recursos 02 analistas de sistemas con 8 horas a la semana.

En la siguiente figura, se presenta el cronograma de actividades:



| Actividad | Semanas | | | | | | | | | | | | | | |
|--|---------|---|---|---|---|---|---|---|---|----|----|----|----|----|----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 1. Entendimiento del negocio | | | | | | | | | | | | | | | |
| 1.1 Determinar los objetivos del negocio | | | | | | | | | | | | | | | |
| 1.2 Evaluar la situación | | | | | | | | | | | | | | | |
| 1.3 Determinar los objetivos de la minería de datos | | | | | | | | | | | | | | | |
| 1.4 Crear un plan para el proyecto de minería de datos | | | | | | | | | | | | | | | |
| 2. Comprensión de los datos | | | | | | | | | | | | | | | |
| 2.1 Recolectar los datos iniciales | | | | | | | | | | | | | | | |
| 2.2 Describir los datos | | | | | | | | | | | | | | | |
| 2.3 Explorar los datos | | | | | | | | | | | | | | | |
| 2.4 Verificar la calidad de los datos | | | | | | | | | | | | | | | |
| 3. Preparación de los datos | | | | | | | | | | | | | | | |
| 3.2 Seleccionar los datos | | | | | | | | | | | | | | | |
| 3.3 Construcción de los datos | | | | | | | | | | | | | | | |
| 3.4 Integrar los datos | | | | | | | | | | | | | | | |
| 3.5 Formatear los datos | | | | | | | | | | | | | | | |
| 4. Modelado | | | | | | | | | | | | | | | |
| 4.1 Seleccionar la técnica de modelado | | | | | | | | | | | | | | | |
| 4.2 Diseñar las pruebas del modelo | | | | | | | | | | | | | | | |
| 4.3 Construir el modelo | | | | | | | | | | | | | | | |
| 4.4 Evaluar el modelo | | | | | | | | | | | | | | | |
| 5. Evaluación | | | | | | | | | | | | | | | |
| 5.1 Evaluar los resultados | | | | | | | | | | | | | | | |
| 5.2 Revisión del proceso | | | | | | | | | | | | | | | |
| 5.3 Determinar las siguientes etapas | | | | | | | | | | | | | | | |

Figura 4.2: Plan del Proyecto.

4.2. Comprensión de los Datos

En esta fase se logró comprender los datos, teniendo presente los objetivos definidos en el punto 4.1.3 se detalla paso a paso lo siguiente:

- Recopilación inicial de datos.
- Descripción de los datos.
- Exploración de los datos.
- Verificación de calidad de datos.

4.2.1. Recolección de Datos Iniciales

La recolección de datos iniciales se realizó a través de fichas de informes operatorios (ver Anexo 1), los cuales forman parte de los libros de cirugía, cada uno en un rango de fechas, contienen datos que fueron escritos a mano y en muchos casos son datos ilegibles.

Estos datos son registrados por el médico encargado de la atención quirúrgica, quien posteriormente entrega la ficha al área de estadística.

En esta etapa se identificaron los siguientes datos:

- Número de informe.
- Fecha de inscripción.
- N° de Historia clínica.
- Datos generales del paciente (Nombres y Apellidos, Edad, Sexo, Ocupación, Procedencia).
- Departamento hospitalario.
- Servicio Hospitalario.

- Datos del servicio hospitalario (N° Piso, N° Cama, N° Sala).
- Días de internamiento antes de la operación.
- Tipo de intervención (Programada, Emergencia).
- Diagnóstico preoperatorio.
- Pre medicación.
- Hora de inicio y Hora Fin de la Anestesia.
- Anestesia.
- Anestésico.
- Cantidad de anestésico.
- Nombre del anestesiólogo.
- Firma del anestesiólogo.
- Descripciones de las intervenciones quirúrgicas.
- Código de las intervenciones quirúrgicas.
- Hora de ingreso y Hora de egreso a la sala de operación.
- N° de la sala de operación.
- Fecha de la operación.
- Hora de inicio y hora de término de la operación.
- Duración de la operación.
- Nombre y Firma del cirujano.
- Nombre y Firma del Primer Ayudante.
- Nombre y Firma del Segundo Ayudante.
- Nombre y Firma del Instrumentista.
- Nombre y Firma del Circulante.
- Protocolo Operatorio.
- Diagnóstico Post Operatorio.

- Hora de ingreso a la sala de operación.
- Hora de salida de la sala de operación.
- Nombre de responsable de la sala de recuperación.
- Complicaciones por anestesia.
- Complicaciones en acto quirúrgico.
- Complicaciones en sala de recuperaciones.
- Condición de egreso (Vivo, Muerto).
- Causa de la condición de ingreso (Anestesia, Operación, Recuperación).
- Suspensión de la condición de egreso.
- Muestra enviada a patología.
- Observaciones.

Por tal razón se desarrolló un sistema de información digitalizando las fichas de informes operatorios que se manejan en físico para la recolección y análisis de los datos.

Para la implementación de dicho sistema se aplicó el Lenguaje de Modelado de Software UML, a continuación se muestran los principales diagramas correspondientes:

a. Arquitectura del Sistema

El sistema se basa en la tecnología cliente servidor y está constituido por 3 componentes principales: navegador web (administrador), servidor web (sistema de cirugías) y base de datos, a continuación se muestra el diagrama de arquitectura:

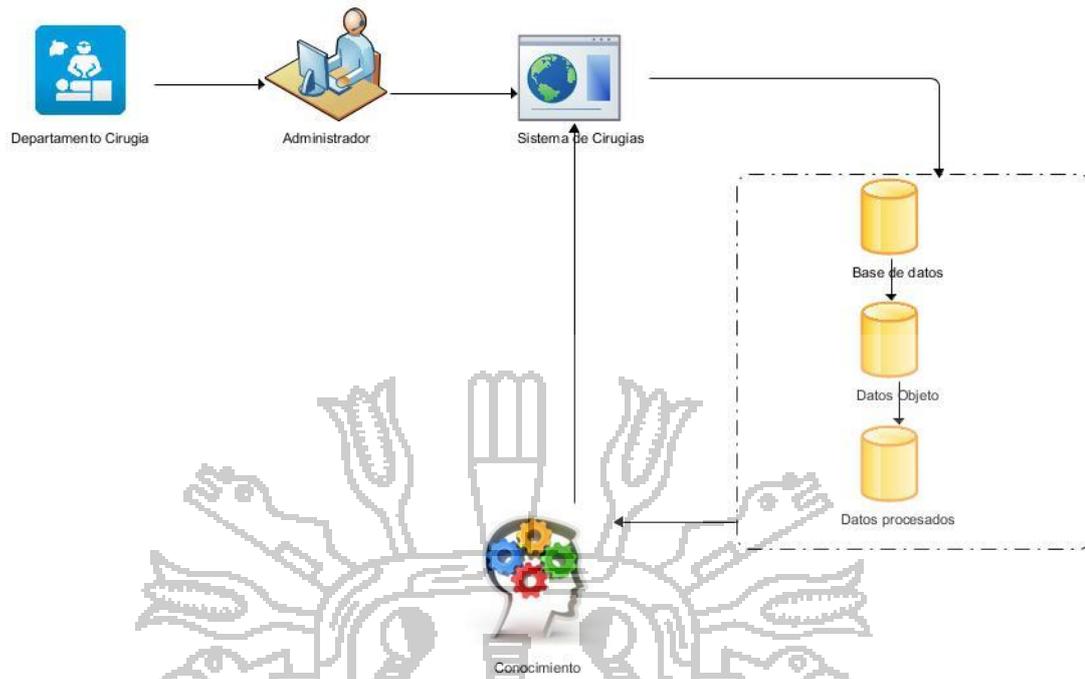


Figura 4.3: Arquitectura de la Herramienta para el Modelo.

b. Casos de Uso

Los actores de los casos de uso son:

- **Actor Administrador:**

Este actor representa al administrador de la herramienta quien alimenta la información en el sistema.

Los casos de usos como resultado de la etapa de análisis de requerimientos para el sistema son:

• **Buscar Paciente:**

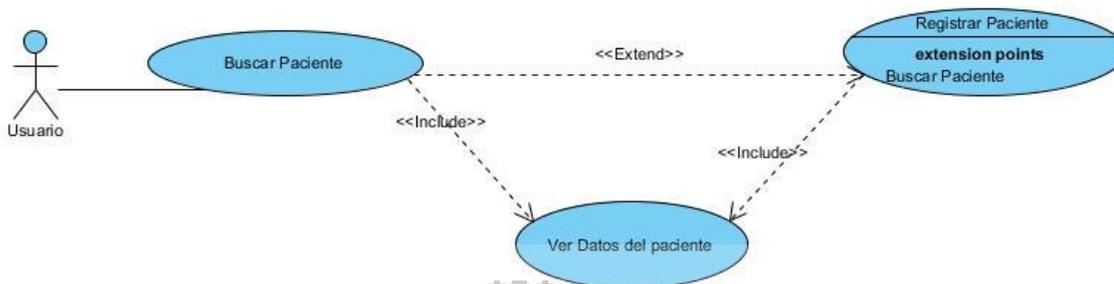


Figura 4.4: Caso de Uso Buscar Paciente (CU-01).

Tabla 4.3: Especificación CU-01.

| CU-01 | Buscar Paciente | |
|------------------|---|--|
| Versión | 1.0 | |
| Actores | Usuario | |
| Descripción | El sistema deberá comportarse tal como se describe en el siguiente caso de uso cuando el usuario ingresa al sistema para buscar un paciente | |
| Precondición | El usuario debe haber iniciado sesión | |
| Secuencia Normal | Paso | Acción |
| | 1. | Usuario ingresa nombre o DNI del paciente |
| | 2. | Sistema verifica y devuelve las coincidencias |
| | 3. | Si la identificación es positiva el sistema muestra: Datos generales del paciente Informe operatorio |
| | 4. | Si la identificación es negativa, sistema muestra registrar paciente |
| | Usuario registra paciente | |
| Pos condición | Ninguna | |
| Excepciones | Paso | Acción |
| | 2. | Si los datos proporcionados no son correctos el sistema muestra mensaje de error |
| Comentarios | Ninguno | |

Fuente: Elaboración Propia.

• Registrar Informe:

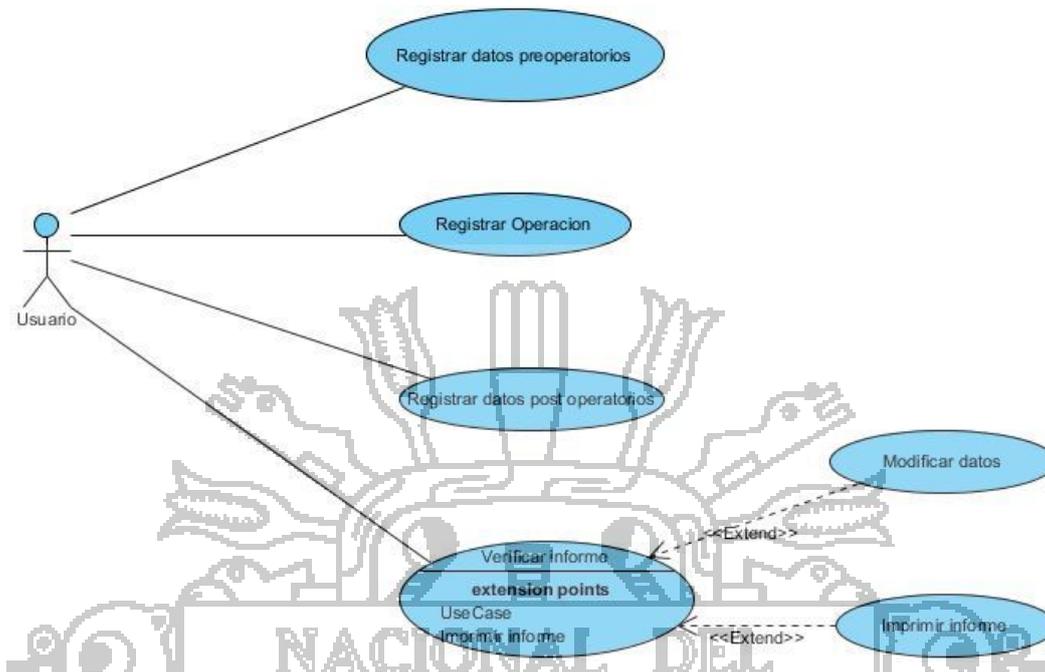


Figura 4.5: Caso de Uso Registrar Informe (CU-02).

Tabla 4.4: Especificación CU-02.

| CU-02 | Registrar informe | |
|------------------|---|--|
| Versión | 1.0 | |
| Actores | Usuario | |
| Descripción | El sistema deberá comportarse tal como se describe en el siguiente caso de uso cuando el usuario desee registrar un informe | |
| Precondición | El usuario debe seleccionar un paciente | |
| Secuencia Normal | Paso | Acción |
| | 1. | Usuario selecciona Nuevo Informe Operatorio |
| | 2. | Sistema devuelve formulario de ingreso |
| | 3. | Usuario ingresa dato del informe |
| | 4. | Usuario selecciona guardar informe |
| | Sistema almacena información de nuevo informe operatorio | |
| Pos condición | Ninguna | |
| Excepciones | Paso | Acción |
| | 3. | Si los datos ingresados no son correctos o ya existen el sistema muestra el mensaje de error |
| Comentarios | Ninguno | |

Fuente: Elaboración Propia.

• Modelar:

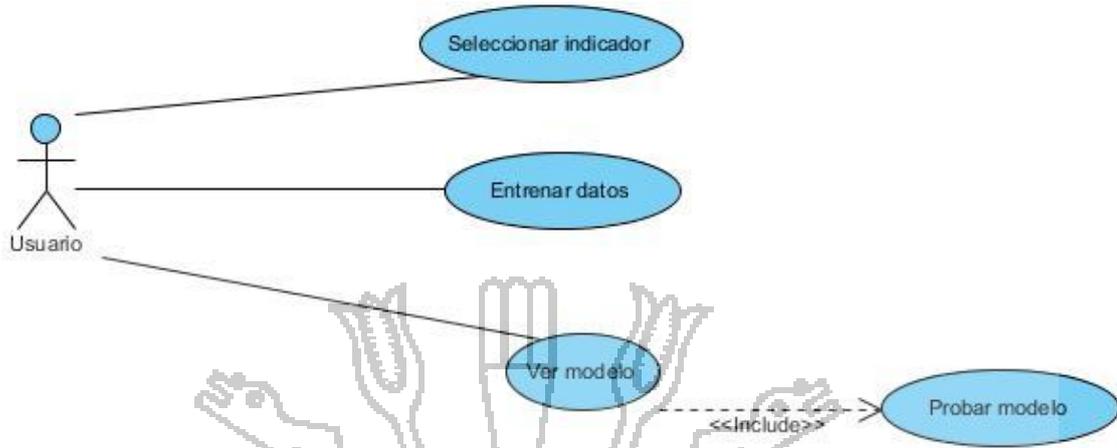


Figura 4.6: Caso de Uso Modelar (CU-03).

Tabla 4.5: Especificación CU-03.

| CU-03 | Modelar | |
|------------------|--|--|
| Versión | 1.0 | |
| Actores | Usuario | |
| Descripción | El sistema deberá comportarse tal como se describe en el siguiente caso de uso cuando el usuario desee modelar | |
| Precondición | Ninguna | |
| Secuencia Normal | Paso | Acción |
| | 1. | Usuario selecciona indicador a modelar |
| | 2. | Usuario selecciona algoritmo de predicción |
| | 3. | Sistema procesa datos |
| | 4. | Sistema muestra datos del indicador seleccionado |
| Pos condición | Ninguna | |
| Excepciones | Paso | Acción |
| | 3. | Si los datos ingresados no son correctos o ya existen el sistema muestra el mensaje de error |
| Comentarios | Ninguno | |

Fuente: Elaboración Propia.

- Ver Estadísticas:

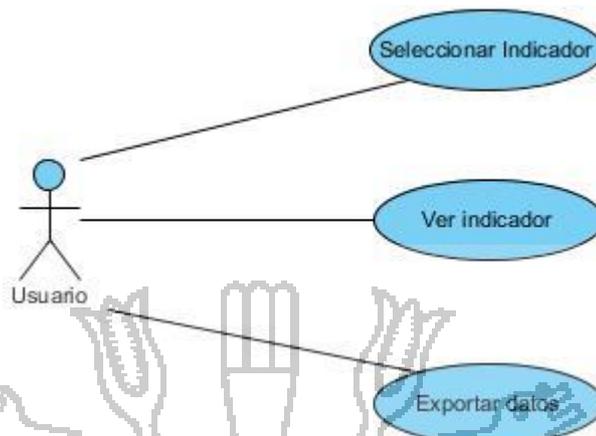


Figura 4.7: Caso de Uso Ver Estadísticas.

Tabla 4.6: Especificación CU-04.

| CU-04 | Ver estadística | |
|------------------|--|--|
| Versión | 1.0 | |
| Actores | Usuario | |
| Descripción | El sistema deberá comportarse tal como se describe en el siguiente caso de uso ver estadística | |
| Precondición | Ninguna | |
| Secuencia Normal | Paso | Acción |
| | 1. | Usuario selecciona indicador |
| | 2. | Sistema devuelve datos del indicador seleccionado por años |
| Pos condición | Ninguna | |
| Excepciones | Paso | Acción |
| | 3. | Si los datos ingresados no son correctos o ya existen el sistema muestra el mensaje de error |
| Comentarios | Ninguno | |

Fuente: Elaboración Propia.

c. Diagramas de Secuencia

Se desarrollaron 4 diagramas de secuencia correspondientes a los casos de uso: buscar paciente, registrar informe, modelar, ver estadísticas, los cuales se detalla a continuación:

Buscar Paciente: Muestra el flujo que se genera entre el usuario y el sistema para realizar la acción registrar informe.

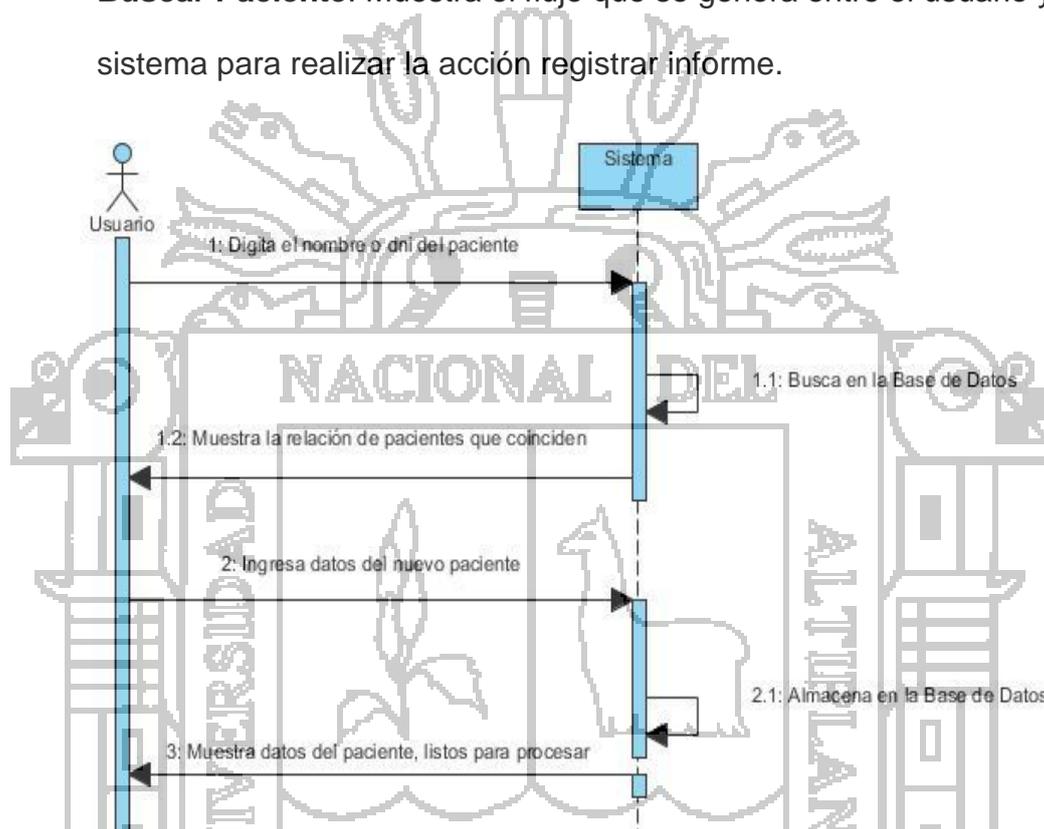


Figura 4.8: Diagrama de Secuencia Buscar Paciente.

Registrar Informe: Se muestra el flujo que se genera entre el usuario y el sistema para realizar la acción de registrar informe.

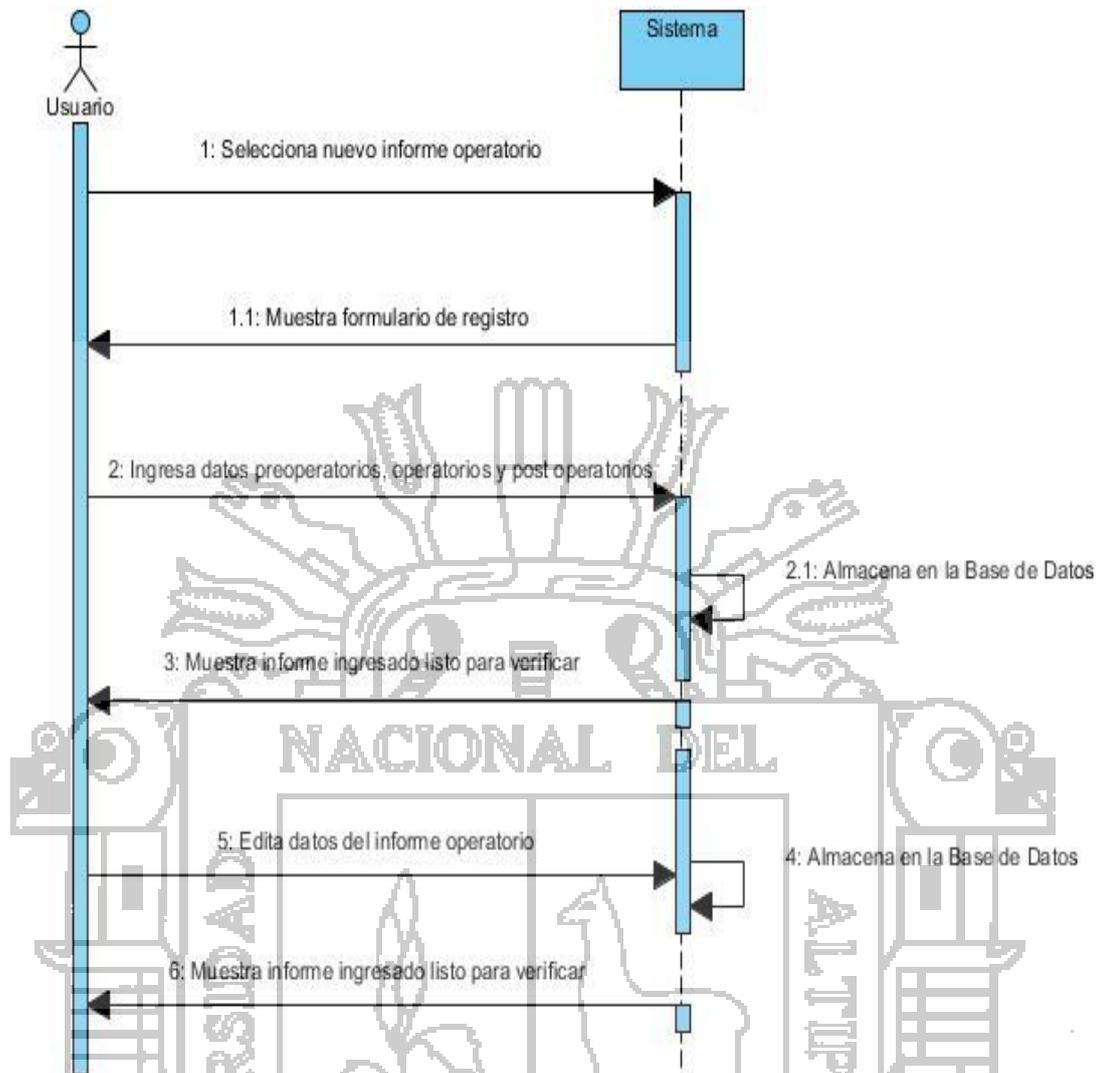


Figura 4.9: Caso de Uso Registrar Informe

Modelar: Muestra el flujo que se genera entre el usuario y el sistema para realizar la acción modelar.

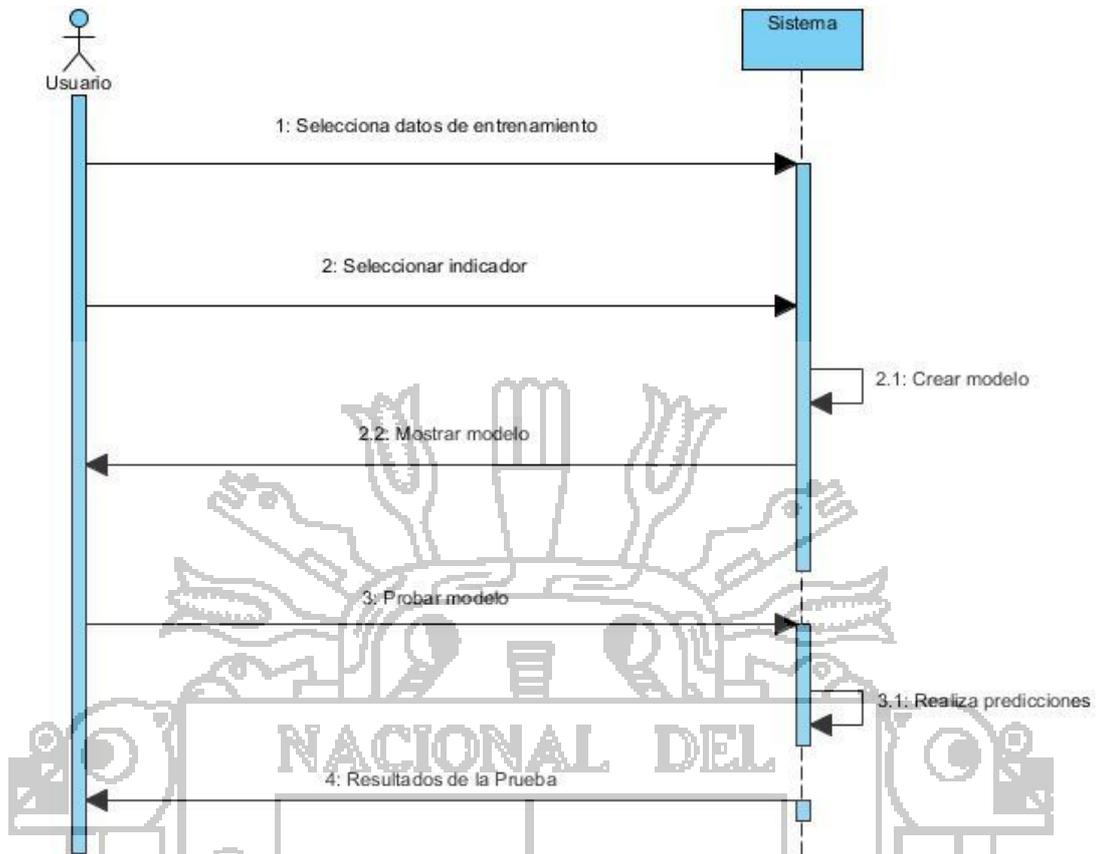


Figura 4.10: Diagrama de Secuencia Modelar.

Ver Estadística: Muestra el flujo que se genera entre el usuario y el sistema para realizar la acción ver estadística.

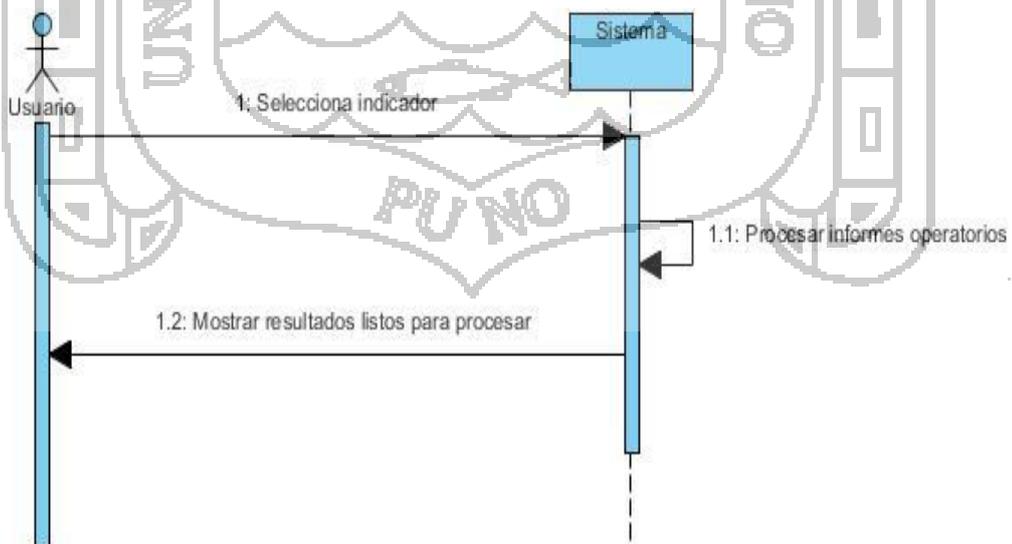


Figura 4.11: Diagrama de Secuencia Ver Estadística.

d. Diagrama de Actividades

El diagrama de actividades proporcionó una vista del comportamiento centrado en la tarea, ya que su propósito es capturar acciones y sus resultados en términos de cambios de estado y a continuación se describe cada uno:

Buscar Paciente: Muestra la secuencia que se genera desde la actividad buscar paciente, hasta que el sistema muestra los datos del paciente, como se muestra en siguiente figura:

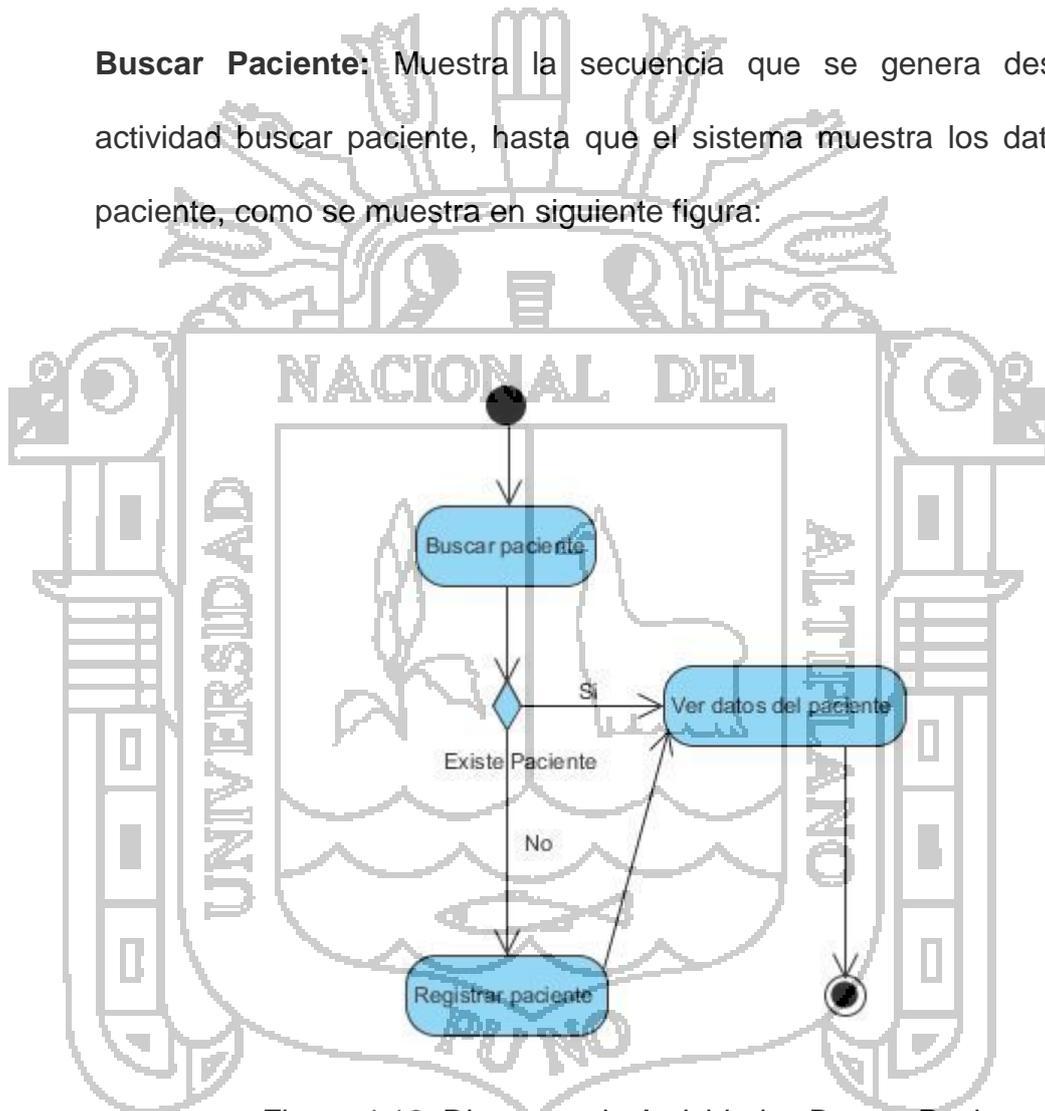


Figura 4.12: Diagrama de Actividades Buscar Paciente.

Registrar informe: Muestra la secuencia que se genera desde la actividad registrar informe operatorio, hasta verificar informe, como se muestra en la siguiente figura:

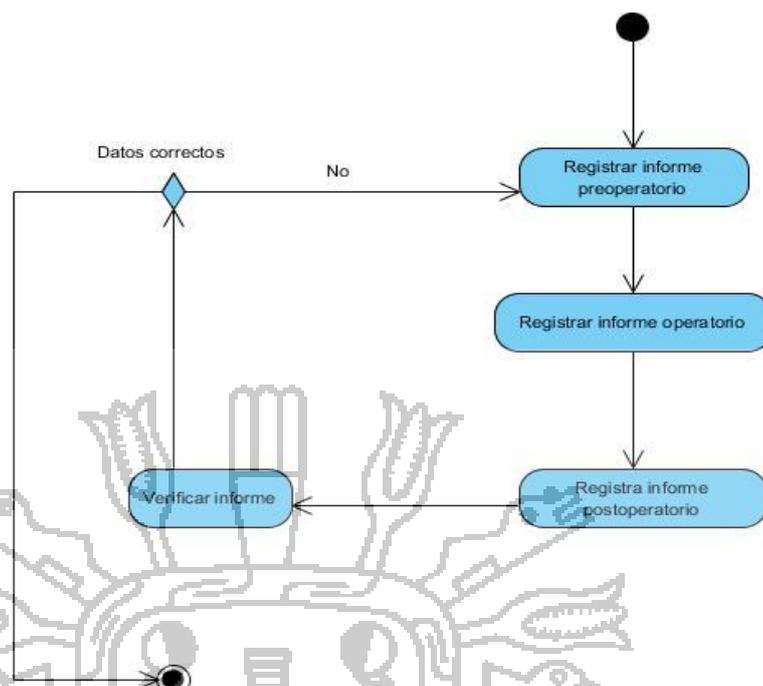


Figura 4.13: Diagrama de Actividades Buscar Paciente.

Modelar: Se detallan las actividades de seleccionar datos de entrenamiento hasta la verificación de modelo, como se muestra en la siguiente figura:

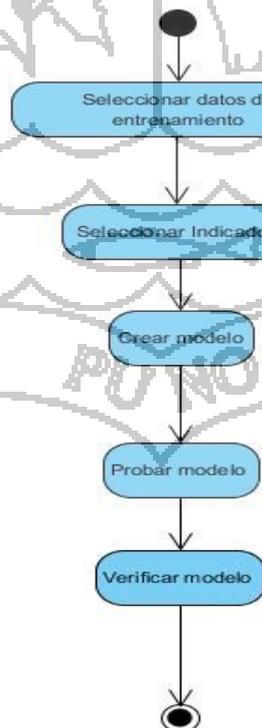


Figura 4.14: Diagrama de Actividades Modelar.

Ver Estadísticas: Se detallan las actividades desde seleccionar indicador hasta la verificación de resultados, como se muestra en la siguiente figura:

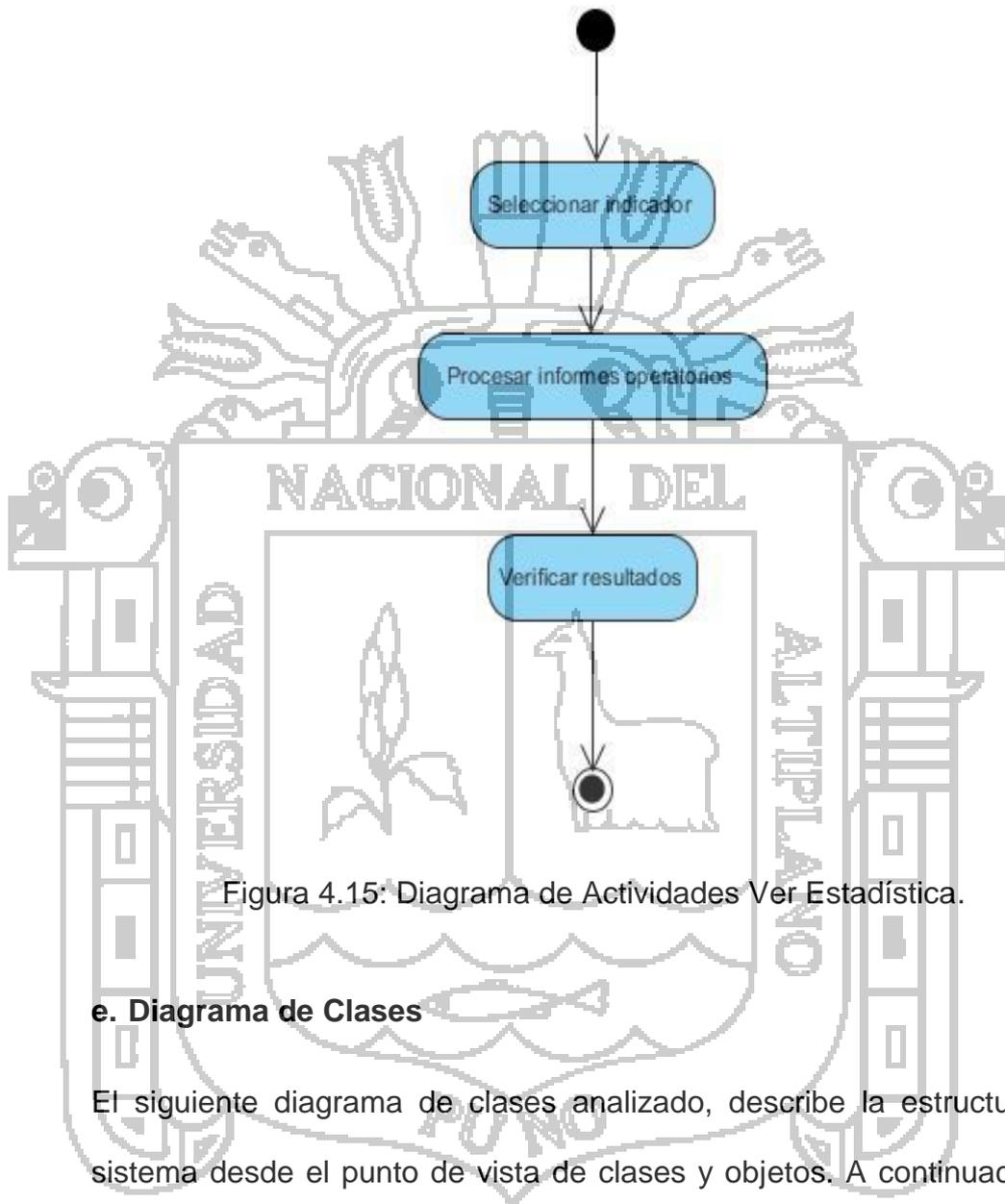


Figura 4.15: Diagrama de Actividades Ver Estadística.

e. Diagrama de Clases

El siguiente diagrama de clases analizado, describe la estructura del sistema desde el punto de vista de clases y objetos. A continuación se muestran las clases que se han identificado y las abstracciones genéricas de sus atributos y comportamiento del conjunto de objetos:

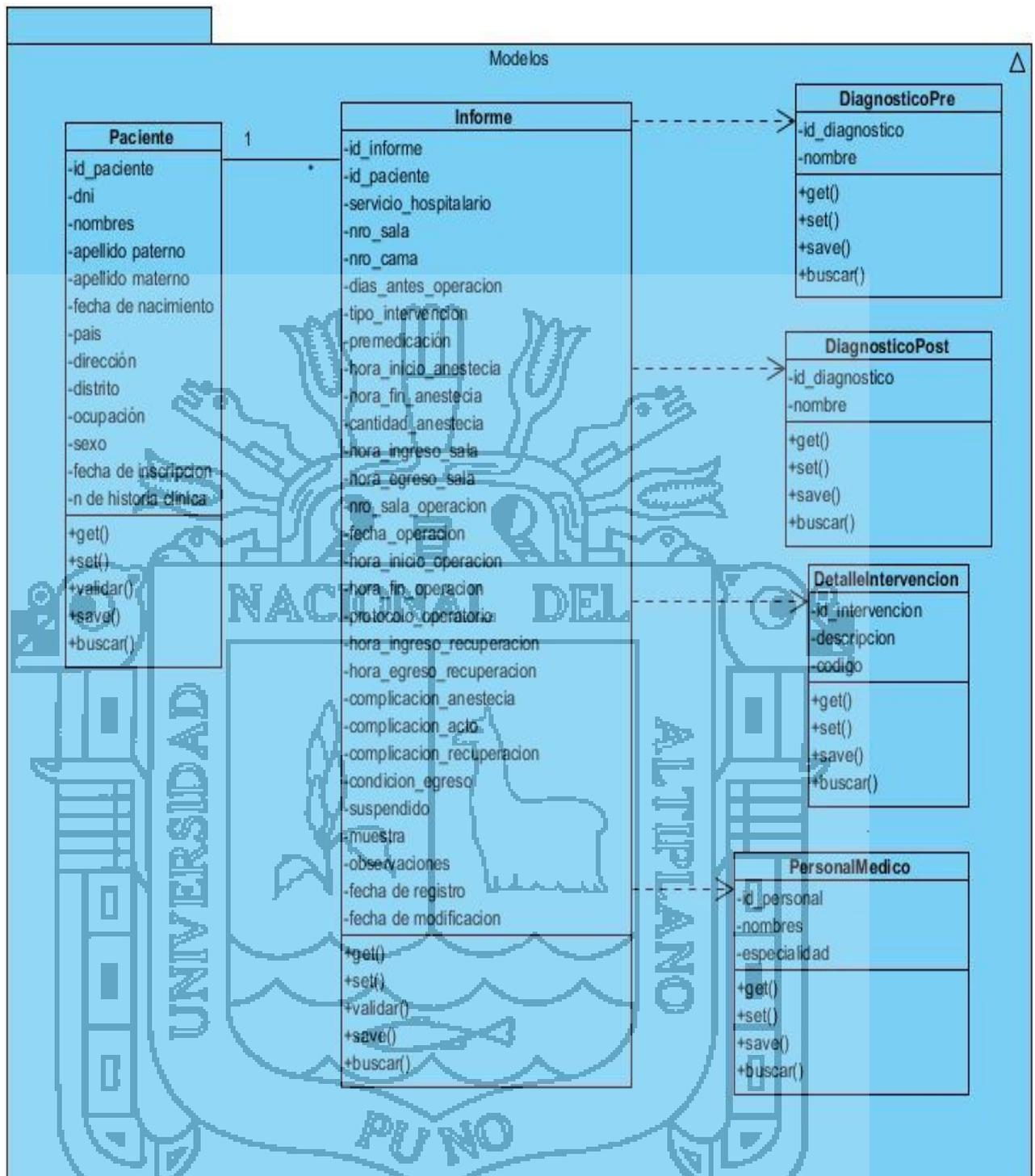


Figura 4.16: Diagrama de Clases.

f. Interfaces Principales de la Herramienta

Se muestran las principales interfaces de la herramienta: En la siguiente figura, se muestra la interfaz de buscar paciente donde el usuario puede hacer la búsqueda a través de DNI, Nombre o Apellidos:

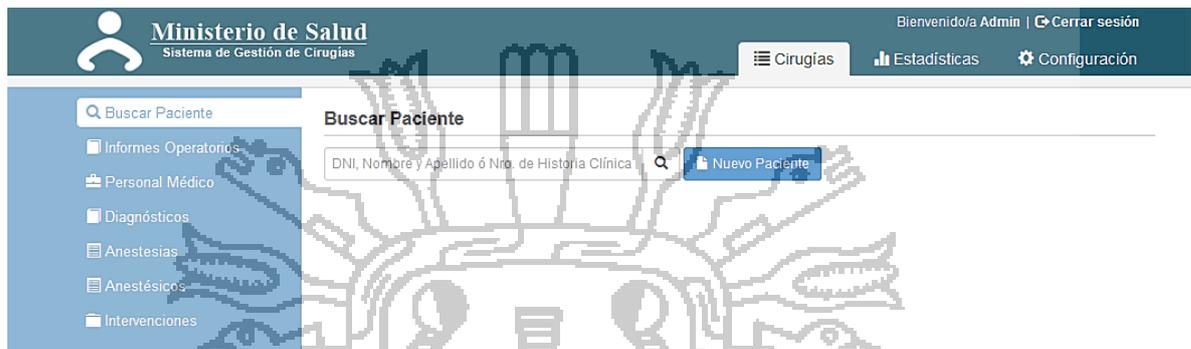


Figura 4.17: Interfaz de Buscar Paciente.

En la siguiente figura se muestra el formulario para el registro de un nuevo paciente, donde es necesario ingresar todos los campos solicitados:

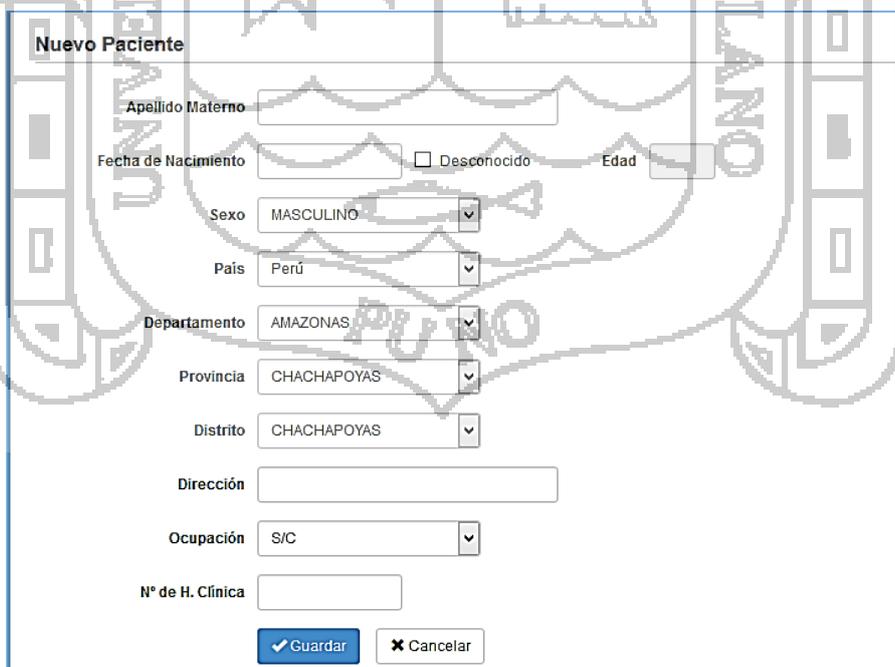


Figura 4.18: Interfaz de Nuevo Paciente.

La interfaz Modificar Nuevo Informe Operatorio muestra el formulario con los campos solicitados para el registro de un nuevo informe operatorio.

Modificar Informe Operatorio FRILAN SULLCA MENDOZA

Paso 1: Datos Pre Operatorios Paso 2: Operación Paso 3: Datos Post Operatorios

Departamento Hospitalario: CIRUGIA

Servicio Hospitalario: CIRUGIA A Sala N°: 3 Cama N°: 14

Días antes de la operación: 3 Intervención: Programada de Emergencia

Diagnóstico Pre Operatorio: Nuevo...

Pre medicación:

Hora Inicio Anestecia: DE 11:35 A

Anestecia: Nuevo...

Anestésico: Nuevo...

Cantidad: 0 %

Anestesiólogo Nombre: Nuevo...

Intervenciones: Nuevo...

Hora de ingreso: 11:10 Sala de Operación N°: 3

Hora de egreso:

Fecha de Operación: 09/05/2013

Hora de inicio: 11:58 Hora de término: 00:00

Cirujano Nombre: Nuevo...

Primer ayudante: - Buscar - Nuevo...

Segundo ayudante: - Buscar - Nuevo...

Instrumentista nombre: Nuevo...

Circulante nombre: Nuevo...

Protocolo Operatorio:

« Anterior Siguiente »

Figura 4.19: Interfaz de Modificar Informe Operatorio - A.

Figura 4.20: Interfaz de Modificar Informe Operatorio – B.

4.2.2. Descripción de los Datos

Se elaboró el siguiente modelo entidad relación que se muestra en la siguiente figura.

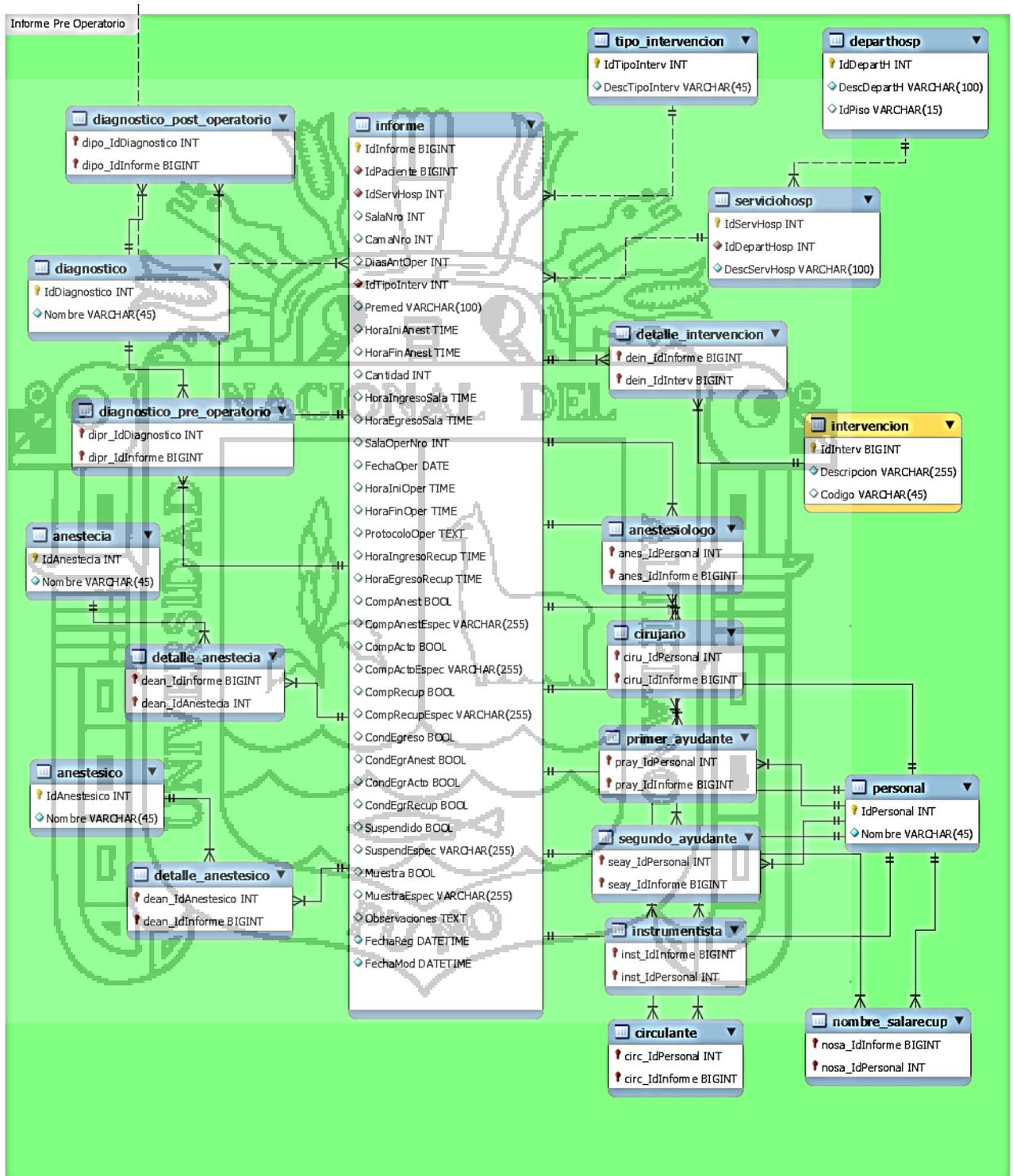


Figura 4.21: Modelo Entidad Relación.

Del sistema se identificó como tabla principal Informe quien almacena la información de los informes operatorios, en la siguiente tabla se describe su estructura.

Tabla 4.7: Descripción de Tabla Informe Operatorios.

| Campo | Descripción |
|------------------|---|
| IdInforme | Código identifica torio del informe operatorio, registra el número de informa |
| IdPaciente | Código identifica torio del paciente, registra el código de identificación del paciente |
| IdServHosp | Código de servicio hospitalario, |
| SalaNro | Numero de sala de operación, donde se encuentra el paciente previo a la operación |
| CamaNro | Numero de cama en la que se realizó la operación |
| DiasAntOper | Cantidad de días antes de la operación |
| IdTipoInterv | Código de identificación de las intervenciones |
| Premed | Nombre de la pre medicación antes de la operación |
| HoraIniAnest | Hora de inicio de anestesia, registra la hora |
| HoraFinAnest | Hora fin de anestesis, registra hora |
| Cantidad | Cantidad de la anestesia, en qué medida se aplica la anestesia |
| HoraIngresoSala | Hora de ingreso a la sala de operación |
| HoraEgresoSala | Hora de salida de la sala de operación |
| SalaOperNro | Numero de sala de operación en la que se interviene al paciente |
| FechaOper | Fecha de operación |
| HoraIniOper | Hora inicio de la operación |
| HoraFinOper | Hora finalización operación |
| ProtocoloOper | Se registra la descripción de todo el protocolo operatorio dentro de la operación |
| HoraIngresoRecup | Hora ingreso a la sala de recuperaciones |
| HoraEgreoRecup | Hora de salida de la sala de operaciones |
| ComAnestEspec | Registra si el paciente sufrió o no complicaciones por anestesia |
| CompActoEspec | Detalla si el paciente sufrió complicaciones en el acto quirúrgico |
| CompRecup | Detalla Complicaciones en salas de recuperaciones en la sala de recuperaciones |

Fuente: Elaboración Propia.

Tabla 4.8: Descripción de Tabla Informe Operatorios.

| | |
|---------------|--|
| ComRecupEspec | Detalla si el paciente sufrió complicaciones en sala de recuperaciones |
| CondEgreso | Condición de egreso vivo, o muerto |
| CondEgrAnest | Condición de egreso por anestesia |
| CondEgrActo | Condición de egreso en acto quirúrgico |
| CondEgrRecup | Condición de egreso de recuperación |
| Suspendido | Estado de la operación |
| SuspendEspec | Detalle de operaciones suspendidas |
| Muestra | Muestras del paciente enviada a patología |
| MuestraEspec | Especificación de la muestra |
| Observaciones | Detalles de la cirugía |
| FechaReg | Fecha de registro |
| FechaMod | Fecha de modificación |

Fuente: Elaboración Propia.

Se trabajó con los datos del periodo 2012, 2013, con un total de 37,800 registros aproximadamente.

4.2.3. Exploración de Datos

Para la exploración de los datos se utilizaron las herramientas de MySQL para el filtrado de los datos.

Para las opciones de procesamiento se utilizó la herramienta WEKA, obteniendo los datos como se muestra en la siguiente figura.

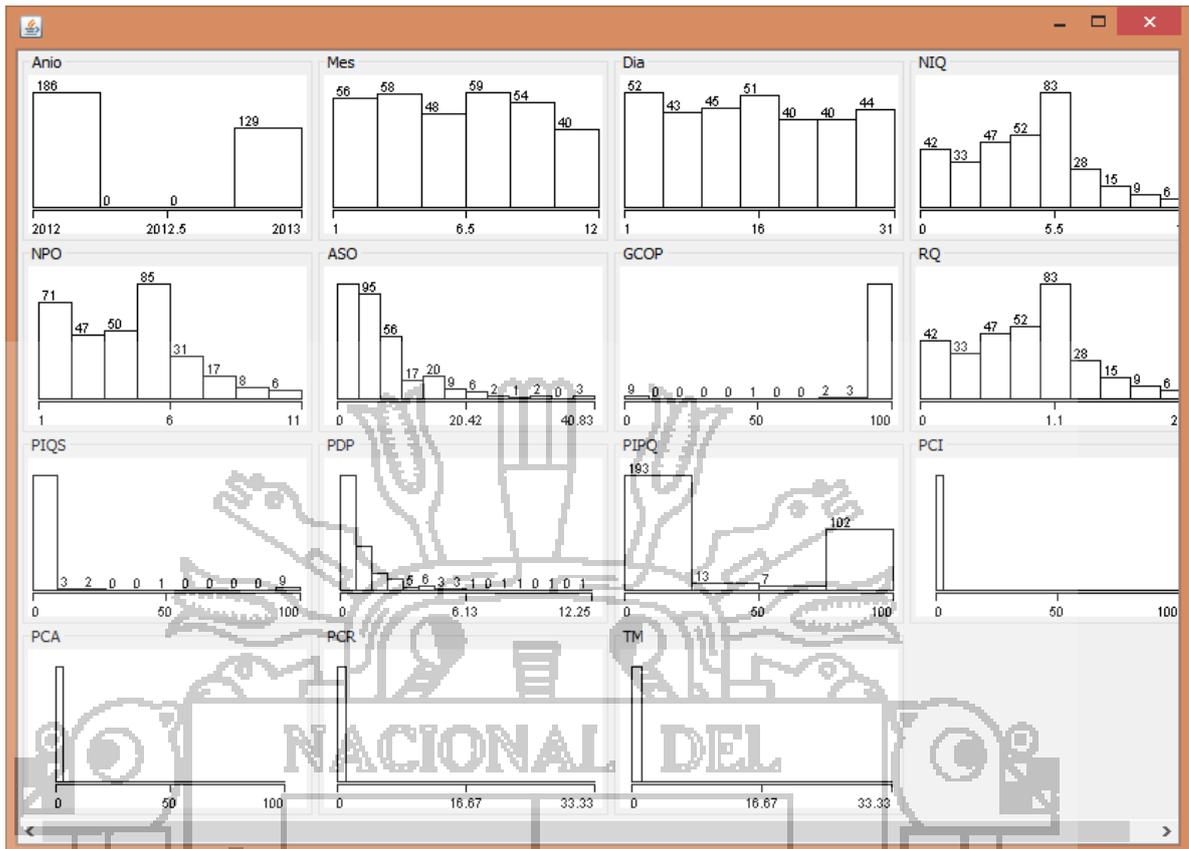


Figura 4.22: Exploración de datos.

4.2.4. Calidad de los Datos

La etapa de verificación de la calidad se realizó de forma muy estrecha con las tareas de selección y limpieza de los datos en la siguiente etapa procesamiento de los datos, dado que las actividades ejecutadas garantizan la consistencia de los datos individuales.

Con la comprensión de los datos se llegó a la conclusión de la siguiente Hipótesis:

Para el criterio de éxito en la mejora para la toma de decisiones del departamento de cirugía el grado de precisión debe ser mayor o igual al 75%.

4.3. Procesamiento de los Datos

En esta fase se procesó los datos a través de los siguientes pasos:

- Selección y limpieza de los datos.
- Estructuración e integración de los datos.
- Formateo de los datos.

4.3.1. Selección y Limpieza de Datos

Los datos seleccionados fueron los requeridos en las fórmulas de los indicadores hospitalarios (ver sustento teórico), los cuales son:

- Fecha de la operación.
- Operación suspendida.
- Id. del paciente operado.
- Hora de ingreso a la sala de operación.
- Hora de término de la operación.
- Horas disponibles de las salas de operaciones (adicional).
- Tipo de intervención.
- Número de días de internamiento antes de la operación.
- Muestra enviada a patología.
- Complicación en acto quirúrgico.
- Complicación por anestesia.
- Complicación en sala de recuperación.
- Condición de egreso.

Para la limpieza de los datos se han identificado 103 registros de informes operatorios en los cuales al menos un dato seleccionado tiene un valor nulo. La siguiente figura muestra la consulta SQL para obtener la cantidad de registros a limpiar.

```
SELECT count(*) FROM `informe`  
WHERE `FechaOper` is null OR FechaOper = '0000-00-00'  
OR `Suspendido` is null  
OR `HoraIngresoSala` is null  
OR `HoraFinOper` is null  
OR `IdTipoInterv` IS NULL  
OR `DiasAntOper` IS NULL  
OR `Muestra` IS NULL  
OR `CompActo` IS NULL  
OR `CompAnest` IS NULL  
OR `CompRecup` IS NULL  
OR `CondEgreso` IS NULL
```

Figura 4.23: Consulta de Registros con Dato Nulo.

Los registros que cumplen la condición de filtrado mediante la consulta mostrada fueron excluidos del archivo de datos objeto de estudio.

4.3.2. Estructuración e Integración de los Datos

Las nuevas estructuras de datos fueron generadas a partir de los indicadores hospitalarios, para lo cual se construyeron consultas SQL en donde fueron integrados los datos seleccionados.

A continuación se muestran las consultas SQL para cada indicador hospitalario con periodicidad mensual:

a. Número de Intervenciones Quirúrgicas (NIQ)

```

SELECT
YEAR(FechaOper) AS Anio
,MONTH(FechaOper) AS Mes
,SUM(IF(Suspendido,0,1)) AS Intervenciones FROM informe
WHERE YEAR(FechaOper)<>0 AND MONTH(FechaOper)<>0
GROUP BY YEAR(FechaOper), MONTH(FechaOper)

```

Figura 4.24: Consulta NIQ.

b. Número de Pacientes Operados (NPO)

```

SELECT
YEAR(FechaOper) AS Anio
,MONTH(FechaOper) AS Mes
,COUNT(DISTINCT IdPaciente) AS Pacientes
FROM informe
WHERE YEAR(FechaOper)<>0 AND MONTH(FechaOper)<>0
GROUP BY YEAR(FechaOper), MONTH(FechaOper)

```

Figura 4.25: Consulta NPO.

c. Aprovechamiento de Salas de Operaciones (ASO)

```

SELECT
YEAR(FechaOper) AS Anio
,MONTH(FechaOper) AS Mes
, HOUR (SEC_TO_TIME (SUM (TIME_TO_SEC (TIMEDIFF (IF (TIMEDIFF (Hora
FinOper, HoraIngresoSala) < 0, ADDTIME (HoraFinOper, '24:00:00'), HoraF
inOper), HoraIngresoSala)))))) / (24 * 30) AS suma FROM informe
WHERE HoraFinOper <> '00:00:00' AND HoraIngresoSala <> '00:00:00'

GROUP BY YEAR(FechaOper), MONTH(FechaOper)

```

Figura 4.26 Consulta ASO.

d. Grado de Cumplimiento de las Operaciones Programadas (GCOP)

```

SELECT
    YEAR(FechaOper) AS Anio
    ,MONTH(FechaOper) AS Mes
    ,COUNT(*) AS Programadas
    ,SUM(CASE WHEN Suspendido > 0
    THEN 1 ELSE 0 END) AS Suspendidos
    ,ROUND(100*((COUNT(*) - SUM(CASE WHEN
    Suspendido > 0 THEN 1
    ELSE 0 END))/COUNT(*)),2) AS Grado
FROM informe
WHERE YEAR(FechaOper)<>0 AND MONTH(FechaOper)<>0
GROUP BY YEAR(FechaOper), MONTH(FechaOper)

```

Figura 4.27: Consulta GCOP.

e. Rendimiento Quirófano (RQ)

```

SELECT
    YEAR(FechaOper) AS Anio
    ,MONTH(FechaOper) AS Mes
    ,ROUND((COUNT(*) - SUM(CASE WHEN Suspendido > 0
    THEN 1 ELSE 0 END))/5,2)
AS Promedio
FROM informe
WHERE YEAR(FechaOper)<>0 AND
MONTH(FechaOper)<>0
GROUP BY YEAR(FechaOper),
MONTH(FechaOper)

```

Figura 4.28: Consulta RQ.

f. Porcentaje de Intervenciones Quirúrgicas Suspendidas (PIQS)

```

SELECT
    YEAR(FechaOper) AS Anio
    ,MONTH(FechaOper) AS Mes
    ,ROUND(100*(SUM(CASE WHEN Suspendido > 0 THEN 1
ELSE 0 END)/COUNT(*)),2) AS Porcentaje
FROM informe WHERE YEAR(FechaOper)<>0
AND MONTH(FechaOper)<>0
GROUP BY YEAR(FechaOper), MONTH(FechaOper)

```

Figura 4.29: Consulta PIQS.

g. Promedio de Días Preoperatorios (PDP)

```

SELECT
    YEAR(FechaOper) AS Anio
    ,MONTH(FechaOper) AS Mes
    ,ROUND(SUM(IF(Suspendido,0,DiasAntOper))
/SUM(IF(Suspendido,0,1)),2) AS Promedio FROM
informe
WHERE YEAR(FechaOper)<>0 AND MONTH(FechaOper)<>0
GROUP BY YEAR(FechaOper), MONTH(FechaOper)

```

Figura 4.30: Consulta PDP.

h. Porcentaje de Infecciones Post Quirúrgicas (PIPQ)

```

SELECT
    YEAR(FechaOper) AS Anio,MONTH(FechaOper) AS Mes
    ,IFNULL (
ROUND(100*SUM(IF(Suspendido,0,Muestra))/SUM(IF(Suspendido,0,1)),
3) , 0 ) AS PIPQ FROM informe
WHERE YEAR(FechaOper)<>0 AND MONTH(FechaOper)<>0
GROUP BY YEAR(FechaOper), MONTH(FechaOper)

```

Figura 4.31: Consulta PIPQ

i. Porcentaje de Complicaciones Intraoperatorias (PCI)

```

SELECT
    YEAR (FechaOper) AS Anio
    ,MONTH (FechaOper) AS Mes
    ,ROUND (SUM (IF (Suspendido,0,CompActo)) /SUM (IF (Suspendido,0,1)),2)
AS Promedio
FROM informe
WHERE YEAR (FechaOper) <>0 AND MONTH (FechaOper) <>0
GROUP BY YEAR (FechaOper), MONTH (FechaOper)

```

Figura 4.32: Consulta PCI.

j. Porcentaje de Complicaciones Anestésicas (PCA)

```

SELECT
    YEAR (FechaOper) AS Anio
    ,MONTH (FechaOper) AS Mes
    ,ROUND (SUM (IF (Suspendido,0,CompAnest)) /SUM (IF (Suspendido,0,1)),2)
) AS Promedio
FROM informe
WHERE YEAR (FechaOper) <>0 AND MONTH (FechaOper) <>0
GROUP BY YEAR (FechaOper), MONTH (FechaOper)

```

Figura 4.33: Consulta PCA.

k. Porcentaje de Complicaciones en Recuperación (PCR)

```

SELECT
    YEAR (FechaOper) AS Anio ,MONTH (FechaOper) AS Mes
    ,ROUND (SUM (IF (Suspendido,0,CompRecup)) /SUM (IF (Suspendido,0,1)),2)
) AS Promedio FROM informe
WHERE YEAR (FechaOper) <>0 AND MONTH (FechaOper) <>0
GROUP BY YEAR (FechaOper), MONTH (FechaOper)

```

Figura 4.34: Consulta PCR.

I. Tasa de Mortalidad (TM)

```
SELECT
    YEAR(FechaOper) AS Anio
    ,MONTH(FechaOper) AS Mes
    ,ROUND(100*(SUM(IF(CondEgreso,0,1))/SUM(IF(Suspendido,0,1)))
    ,2) AS Promedio
FROM informe
WHERE YEAR(FechaOper)<>0 AND MONTH(FechaOper)<>0
GROUP BY YEAR(FechaOper), MONTH(FechaOper)
```

Figura 4.35: Consulta TM.

4.3.3. Formateo de Datos

Los datos iniciales fueron modificados redondeando a dos cifras los valores finales con el fin de lograr un mayor grado de confiabilidad con el modelo.

Los valores resultantes de los indicadores hospitalarios fueron formateados redondeados a dos cifras decimales, por otro lado las fechas fueron separadas por año, mes y día, siendo convertidos, cada uno de estos, en valores numéricos.

4.3.4. Implementación del Modelado

En esta fase se implementó el modelo a través de los siguientes pasos:

- Selección de la técnica de modelamiento.
- Diseño de las pruebas.
- Construcción del modelo.
- Evaluación técnica.

4.3.5. Selección de Técnica de Modelamiento

En esta etapa del proceso se seleccionó 02 técnicas clasificadas dentro de la categoría supervisados o predictivos, estos son: Regresión Lineal y Árboles de Decisión, con el propósito de cumplir los objetivos de esta investigación. Los algoritmos supervisados o predictivos predicen el valor de un atributo de un conjunto de datos, a partir de datos cuyo atributo se conoce se induce una relación entre dicho atributo y otra serie de atributos.

La regresión lineal se denomina al análisis de la relación entre una variable denominada variable dependiente y un conjunto de otras variables, denominadas variables independientes, cuyo objetivo es predecir el valor de la variable dependiente en función del valor de las variables independientes.

Los árboles de decisión son una forma de representación sencilla, muy usada entre los sistemas de aprendizaje supervisado, se basan en la partición de ejemplos según ciertas condiciones que se aplican a los valores de los atributos.

4.3.6. Diseño de las Pruebas

En este punto se seleccionó el conjunto de entrenamiento y el conjunto de pruebas:

a. Conjunto de entrenamiento

Para la construcción del conjunto de entrenamiento se utilizó la herramienta de filtros no supervisados aplicados a conjunto de datos de

WEKA. Al conjunto de datos iniciales, se le aplicó un filtro de instancias no supervisado, cuyo formato es el siguiente:

```
weka.filters.unsupervised.instance.RemovePercentage
```

Figura 4.36: Formato del Filtro de la Instancia RemovePercentage.

El resultado que se obtiene en este caso es eliminar del conjunto de datos el número de casos correspondiente al porcentaje suministrado como parámetro.

Una vez aplicado el filtro se obtuvieron los conjuntos de entrenamiento los cuales se muestran en la siguiente tabla:

Tabla 4.9: Diseño de Pruebas.

| Prueba | Conjunto de validación : 315 instancias | |
|--------|---|--------|
| | Conjuntos de entrenamiento | |
| | Tamaño | Nombre |
| 1 | 158 | TM |
| 2 | 158 | NIQ |
| 3 | 158 | NPO |
| 4 | 158 | ASO |
| 5 | 158 | GCOP |
| 6 | 158 | RQ |
| 7 | 158 | PIPQS |
| 8 | 158 | PDP |
| 9 | 158 | PIPQ |
| 10 | 158 | PCI |
| 11 | 158 | PCA |
| 12 | 158 | PCR |
| 13 | 158 | TM |

Fuente: Elaboración Propia.

4.3.7. Construcción del Modelo

El formato de aplicación del algoritmo es el siguiente:

```
weka.classifiers.functions.LinearRegression
```

Figura 4.37: Formato del algoritmo de Regresión Lineal

A continuación se presenta el modelo y su evaluación en forma detallada por cada indicador hospitalario.

En la tabla 4.10, se detalla los parámetros que mide el modelo de las técnicas seleccionadas que son: Regresión Lineal, Árboles de decisión.

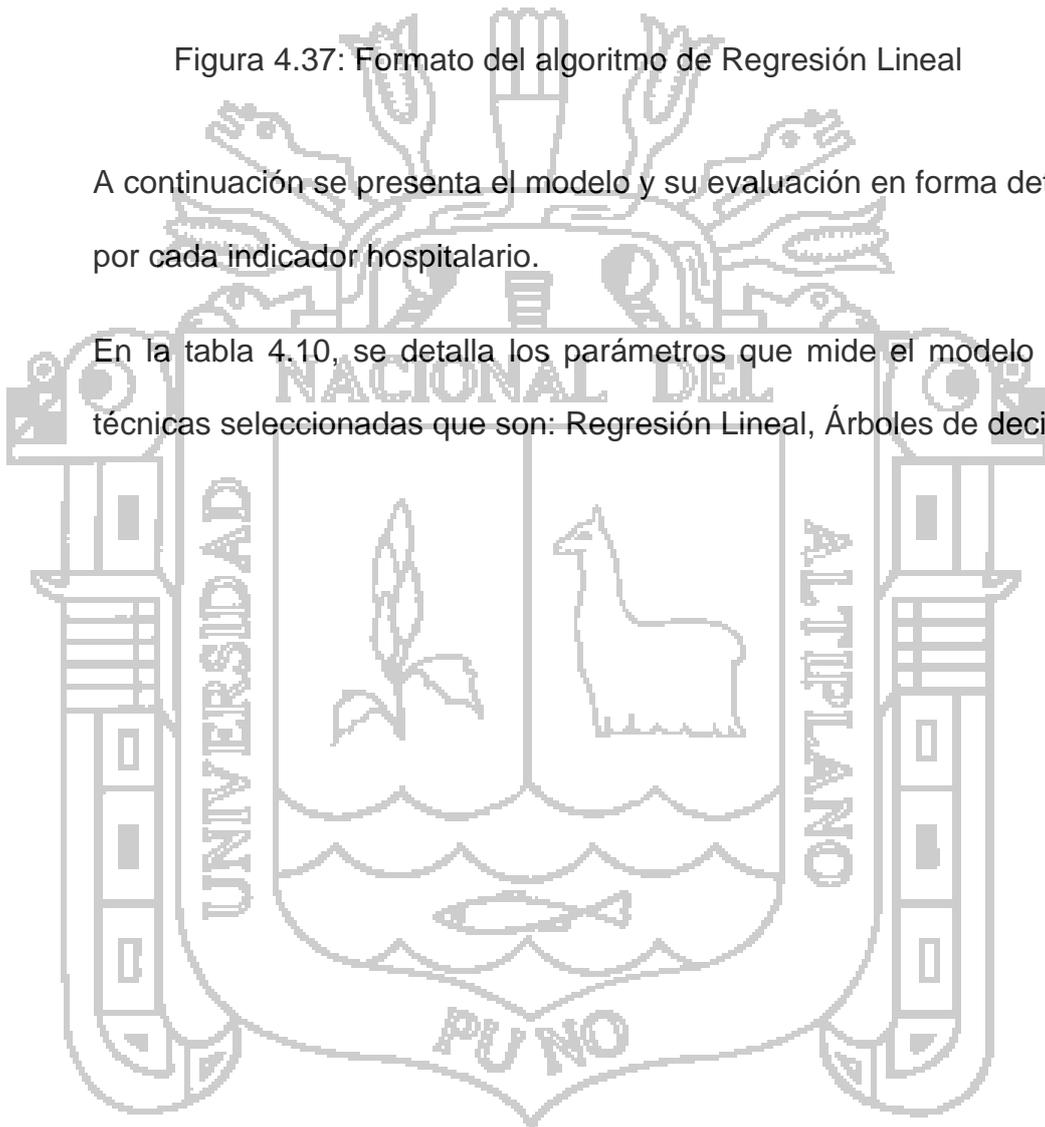


Tabla 4.10: Selección de la Técnica

| Indicadores Hospitalarios | Indicadores de Evaluación | Regresión Lineal | Arboles de decisión |
|--|-----------------------------|------------------|---------------------|
| Número de Intervenciones quirúrgicas | Correlation coefficient | 1 | 0.8272 |
| | Mean absolute error | 0 | 1.1216 |
| | Root mean squared error | 0 | 1.3295 |
| | Relative absolute error | 0.00% | 57.24% |
| | Root relative squared error | 0.00% | 56.01% |
| | Total Number of Instances | 315 | 315 |
| Número de Pacientes Operados | Correlation coefficient | 0.9786 | 0.812 |
| | Mean absolute error | 0.1313 | 1.092 |
| | Root mean squared error | 0.4717 | 1.3375 |
| | Relative absolute error | 6.88% | 57.20% |
| | Root relative squared error | 20.58% | 58.36% |
| | Total Number of Instances | 315 | 315 |
| Aprovechamiento de Salas de Operaciones | Correlation coefficient | 0.9911 | 0.52 |
| | Mean absolute error | 0.1365 | 3.65 |
| | Root mean squared error | 0.8037 | 5.51 |
| | Relative absolute error | 14.16% | 79.80% |
| | Root relative squared error | 13.33 | 85.11% |
| | Total Number of Instances | 315 | 315 |
| Grado de Cumplimiento de Operaciones Programadas | Correlation coefficient | 0.6128 | 0.97 |
| | Mean absolute error | 3.2197 | 0.81 |
| | Root mean squared error | 5.1245 | 3.43 |
| | Relative absolute error | 70.24% | 13.06% |
| | Root relative squared error | 79.02% | 20.24% |
| | Total Number of Instances | 315 | 315 |
| Rendimiento de Quirófano | Correlation coefficient | 1 | 0.82 |
| | Mean absolute error | 0 | 0.22 |
| | Root mean squared error | 0 | 0.26 |
| | Relative absolute error | 0.00% | 57.24% |
| | Root relative squared error | 0.00% | 56.01% |
| | Total Number of Instances | 315 | 315 |

Fuente: Elaboración Propia.

Tabla 4.11: Selección de la Técnica

| Indicadores Hospitalarios | Indicadores de Evaluación | Regresión Lineal | Arboles de decisión |
|--|-----------------------------|------------------|---------------------|
| Promedio de Intervenciones Quirúrgicas Suspendidas | Correlation coefficient | 1 | 0.82 |
| | Mean absolute error | 0 | 0.22 |
| | Root mean squared error | 0 | 0.26 |
| | Relative absolute error | 0.00% | 57.24% |
| | Root relative squared error | 0.00% | 56.01% |
| | Total Number of Instances | 315 | 315 |
| Promedio de Días Preoperatorios | Correlation coefficient | 0.257 | 0.19 |
| | Mean absolute error | 0.9335 | 0.95 |
| | Root mean squared error | 1.4796 | 1.50 |
| | Relative absolute error | 97.25% | 99.04% |
| | Root relative squared error | 96.64% | 98.15% |
| | Total Number of Instances | 315 | 315 |
| Porcentaje de Intervenciones Post Quirúrgicas | Correlation coefficient | 0.9911 | 0.64 |
| | Mean absolute error | 0.1365 | 26.90 |
| | Root mean squared error | 0.8037 | 34.33 |
| | Relative absolute error | 14.16% | 63.88% |
| | Root relative squared error | 13.33 | 76.04% |
| | Total Number of Instances | 315 | 315 |
| Porcentaje de Complicaciones Intraoperatorias | Correlation coefficient | 0.98 | 0.98 |
| | Mean absolute error | 0.13 | 0.10 |
| | Root mean squared error | 0.81 | 0.93 |
| | Relative absolute error | 18.54% | 14.00% |
| | Root relative squared error | 14.24% | 16.44% |
| | Total Number of Instances | 315 | 315 |
| Porcentaje de Complicaciones Anestésicas | Correlation coefficient | 0.99 | 0.93 |
| | Mean absolute error | 0.13 | 0.33 |
| | Root mean squared error | 0.80 | 2.18 |
| | Relative absolute error | 14.16% | 34.89% |
| | Root relative squared error | 13.33% | 36.24% |
| | Total Number of Instances | 315 | 315 |

Fuente: Elaboración Propia.

Tabla 4.12: Selección de la Técnica

| Indicadores Hospitalarios | Indicadores de Evaluación | Regresión Lineal | Árboles de decisión |
|--|-----------------------------|------------------|---------------------|
| Porcentaje de Complicaciones en Recuperación | Correlation coefficient | 0.66 | 0.87 |
| | Mean absolute error | 0.39 | 0.12 |
| | Root mean squared error | 1.64 | 1.12 |
| | Relative absolute error | 117.01% | 37.61% |
| | Root relative squared error | 75.11% | 51.52% |
| | Total Number of Instances | 315 | 315 |
| Tasa de Mortalidad | Correlation coefficient | 0.66 | 0.75 |
| | Mean absolute error | 0.48 | 0.31 |
| | Root mean squared error | 1.84 | 1.61 |
| | Relative absolute error | 94.19% | 59.78% |
| | Root relative squared error | 74.81% | 65.54% |
| | Total Number of Instances | 315 | 315 |

Fuente: Elaboración Propia.

Del cuadro, tenemos que en la mayoría de indicadores a excepción de Porcentaje de complicaciones en recuperación, tasa de mortalidad y Grado de Cumplimiento de Operaciones Programadas la técnica de Regresión Lineal es más efectiva que los Árboles de Decisión.

A continuación se detalla los modelos detallados para cada indicador con Regresión Lineal.

a. Número de Intervenciones Quirúrgicas (NIQ)

La unidad de medida del número de intervenciones quirúrgicas es en unidades y el modelo obtenido para este indicador es el siguiente:

Linear Regression Model

$$\text{NIQ} = 5 * \text{RQ} + 0$$

Time taken to build model: 0 seconds

Figura 4.38: Modelo Regresión Lineal NIQ en Unid.

M5 pruned model tree:

(using smoothed linear models)

LM1 (315/0%)

LM num: 1

$$\text{NIQ} = 5 * \text{RQ} + 0$$

Number of Rules : 1

Time taken to build model: 0.03 seconds

Figura 4.39: Modelo Árboles de Decisión NIQ en Unid.

A continuación se muestra la precisión de la predicción con modelo hasta el julio del 2016 comparado con la situación actual hasta mayo 2013 por tanto se concluye que para NIQ la precisión de la predicción es exacta y depende del RQ:

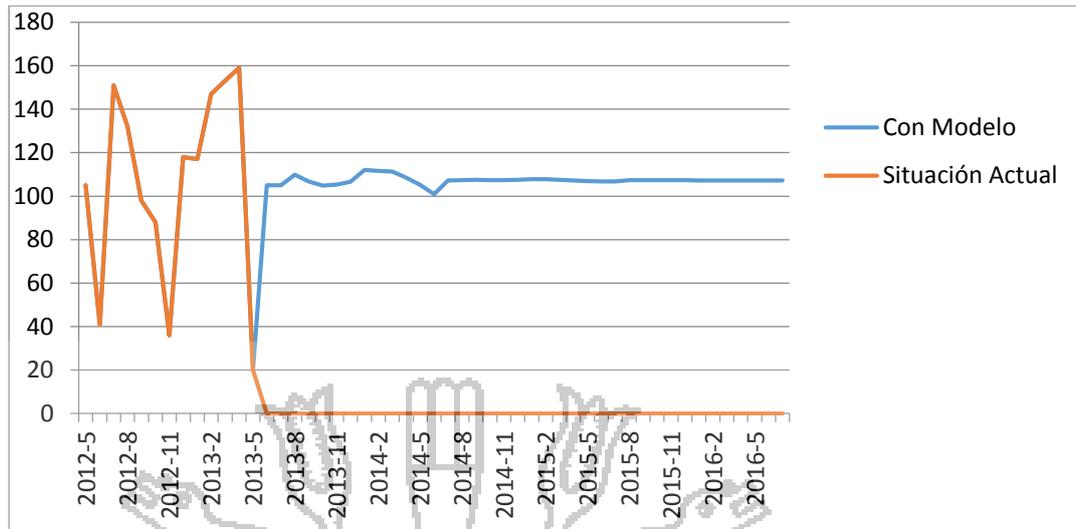


Figura 4.40: Predicción con Regresión Lineal NIQ

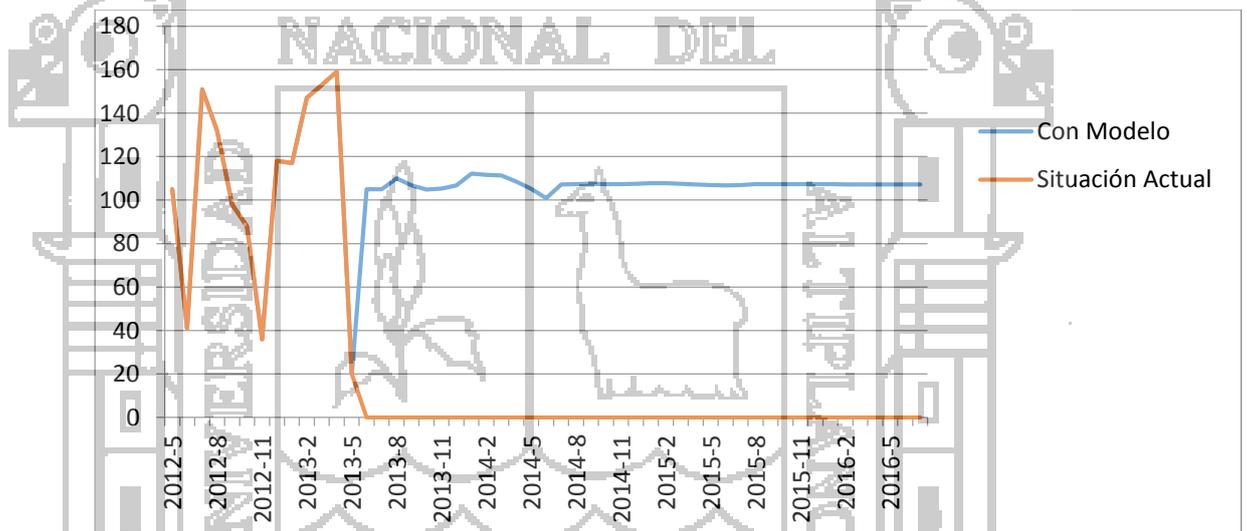


Figura 4.41: Predicción con Árboles de Decisión NIQ.

Donde se puede observar que NIQ está en función de RQ, lo cual es correcto por lo que se obtiene la siguiente evaluación:

```

=== Evaluation on training set ===

=== Summary ===

Correlation coefficient          1
Mean absolute error            0
Root mean squared error        0
Relative absolute error        0    %
Root relative squared error    0    %
Total Number of Instances      315

```

Figura 4.42: Evaluación con Regresión Lineal NIQ.

```

=== Cross-validation ===
=== Summary ===

Correlation coefficient          0.9828
Mean absolute error            0.1421
Root mean squared error        0.4234
Relative absolute error        7.3843 %
Root relative squared error    18.3681 %
Total Number of Instances      315

```

Figura 4.43: Evaluación con Árboles de Decisión NIQ.

Esta prueba crea un modelo el cual obtiene un coeficiente de correlación de 1 con regresión lineal, lo que significa que clasifica correctamente el 100% de instancias.

b. Número de Pacientes Operados (NPO)

La unidad de medida del Número de intervenciones es en Unidades y el modelo obtenido para este indicador es el siguiente:

Linear Regression Model

$$\text{NPO} = 0.4883 * \text{NIQ} + 0.0133 * \text{ASO} - 0.0174 * \text{GCOP} + 2.4414 * \text{RQ} + 0.0174 * \text{PIQS} + 1.7463$$

Time taken to build model: 0 seconds

Figura 4.44: Modelo con Regresión Lineal NPO en Unid.

M5 pruned model tree:

LM num: 1

$$\text{NPO} = 0.2156 * \text{NIQ} + 0.1466 * \text{ASO} - 0.0061 * \text{GCOP} + 1.4055$$

LM num: 2

$$\text{NPO} = -0.055 * \text{Mes} + 0.2156 * \text{NIQ} + 0.2537 * \text{ASO} - 0.0061 * \text{GCOP} + 1.9877$$

LM num: 3

$$\text{NPO} = -0.055 * \text{Mes} + 0.2156 * \text{NIQ} + 0.2537 * \text{ASO} - 0.0061 * \text{GCOP} + 2.0238$$

LM num: 4

$$\text{NPO} = -0.0641 * \text{Mes} + 0.2156 * \text{NIQ} + 0.2609 * \text{ASO} - 0.0061 * \text{GCOP} + 1.8619$$

LM num: 5

$$\text{NPO} = 0.2156 * \text{NIQ} + 0.0917 * \text{ASO} - 0.0061 * \text{GCOP} + 1.9185$$

LM num: 6

$$\text{NPO} = 0.2362 * \text{NIQ} + 0.0455 * \text{ASO} - 0.0061 * \text{GCOP} + 2.669$$

LM num: 7

$$\text{NPO} = 0.2665 * \text{NIQ} + 0.0107 * \text{ASO} - 0.0452 * \text{GCOP} + 7.3879$$

LM num: 8

$$\text{NPO} = 0.4421 * \text{NIQ} + 0.0013 * \text{ASO} - 0.0033 * \text{GCOP} + 3.1374$$

LM num: 9

$$\text{NPO} = 0.4241 * \text{NIQ} + 0.0013 * \text{ASO} - 0.0033 * \text{GCOP} + 3.7894$$

LM num: 10

$$\text{NPO} = 0.4241 * \text{NIQ} + 0.0149 * \text{ASO} - 0.0033 * \text{GCOP} + 3.6923$$

LM num: 11

$$\text{NPO} = 0.4241 * \text{NIQ} + 0.0013 * \text{ASO} - 0.0033 * \text{GCOP} + 3.8059$$

LM num: 12

$$\text{NPO} = 0.4241 * \text{NIQ} + 0.0013 * \text{ASO} - 0.0033 * \text{GCOP} + 3.7957$$

LM num: 13

$$\text{NPO} = 0.9838 * \text{NIQ} + 0.0013 * \text{ASO} - 0.0033 * \text{GCOP} + 0.4307$$

Number of Rules : 13

Time taken to build model: 0.04 seconds

Figura 4.45: Modelo Árboles de Decisión NPO en Unid.

A continuación se muestra el resultado de la predicción hasta el 2016 con modelo comparado con la situación actual con los datos hasta mayo

2013 en el que se concluye que para el NPO la precisión de la predicción es exacta:

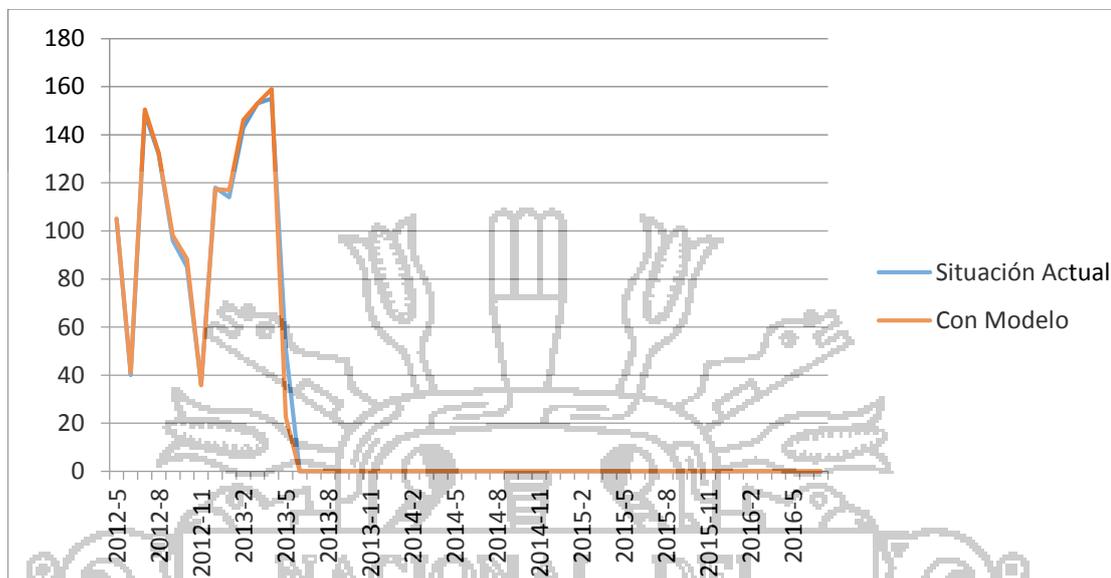


Figura 4.46: Predicción con Regresión Lineal NPO

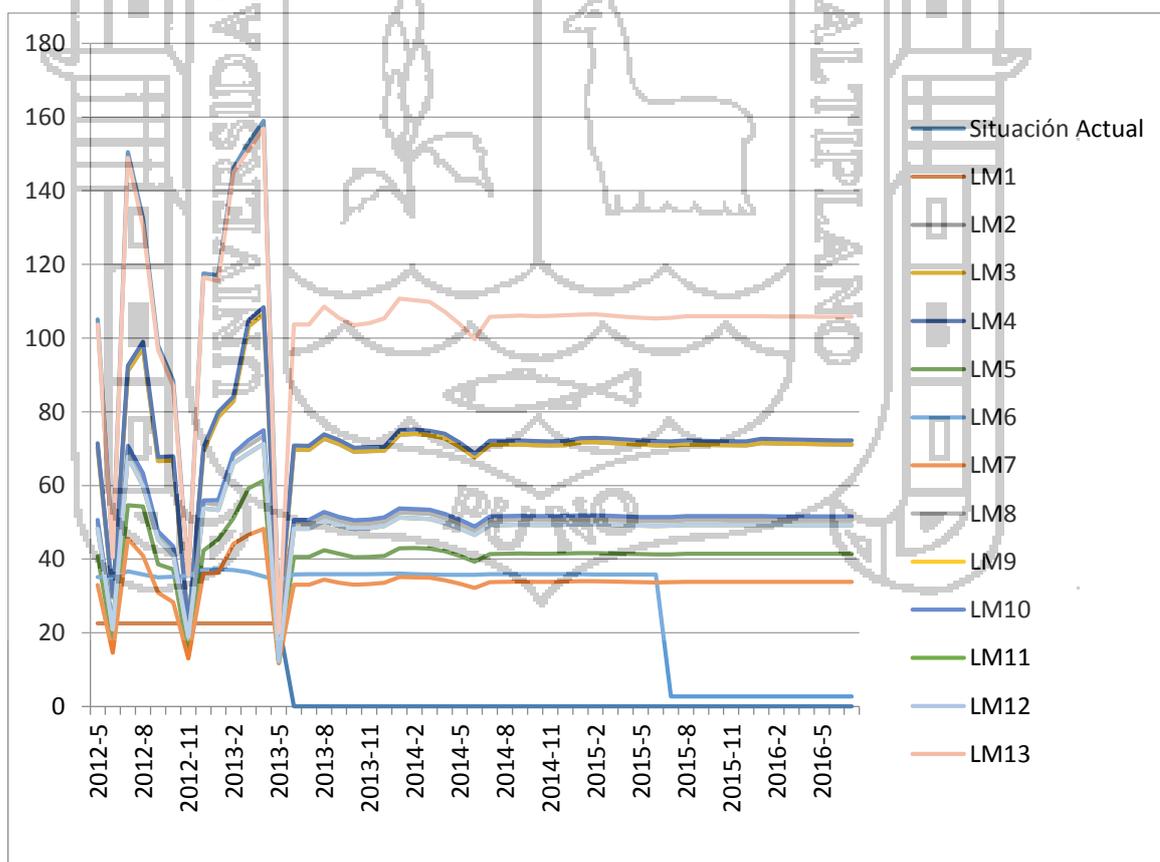


Figura 4.47: Predicción con Árboles de Decisión NPO

Donde se puede observar que NPO está en función de NIQ, ASO, GCOP, RQ y PIQS, con lo cual se obtiene la siguiente evaluación:

```

=== Evaluation on training set ===
=== Summary ===
Correlation coefficient           0.9786
Mean absolute error              0.1313
Root mean squared error          0.4717
Relative absolute error          6.8784 %
Root relative squared error      20.5848 %
Total Number of Instances       315
  
```

Figura 4.48: Evaluación con Regresión Lineal NPO.

```

=== Cross-validation ===
=== Summary ===
Correlation coefficient           0.9828
Mean absolute error              0.1421
Root mean squared error          0.4234
Relative absolute error          7.3843 %
Root relative squared error      18.3681 %
Total Number of Instances       315
  
```

Figura 4.49: Evaluación con Árboles de Decisión NPO.

Esta prueba crea un modelo el cual obtiene un coeficiente de correlación de 0.9786, lo que significa que clasifica correctamente el 98% de instancias con regresión lineal y con árboles de decisión igual.

c. Aprovechamiento de Salas de Operaciones (ASO)

El modelo obtenido para este indicador es el siguiente:

| |
|--|
| <p style="text-align: center;">Linear Regression Model</p> $\text{ASO} = 1.6863 * \text{NPO} + 0.4863 * \text{PDP} + 0.0107 * \text{PIPQ} + 0.5329 * \text{PCI} + -0.4665 * \text{PCA} + 0.4132 * \text{PCR} + -1.3152$ <p style="text-align: center;">Time taken to build model: 0 seconds</p> |
|--|

Figura 4.50: Modelo con Regresión Lineal ASO en Horas.

| |
|---|
| <p style="text-align: center;">M5 pruned model tree:</p> <p>LM num: 1</p> $\text{ASO} = 1.6792 * \text{NPO} + 0.4925 * \text{PDP} + 0.0095 * \text{PIPQ} - 1.2064$ <p style="text-align: center;">Number of Rules : 1</p> <p style="text-align: center;">Time taken to build model: 0.14 seconds</p> |
|---|

Figura 4.51: Modelo Con Árboles de Decisión ASO en Horas

A continuación se muestra el resultado de la predicción hasta julio del 2016 con modelo comparado con la situación actual con los datos hasta mayo del 2013 por lo que se concluye que el ASO hay una ligera variación en la precisión de la predicción y se calcula por hora:

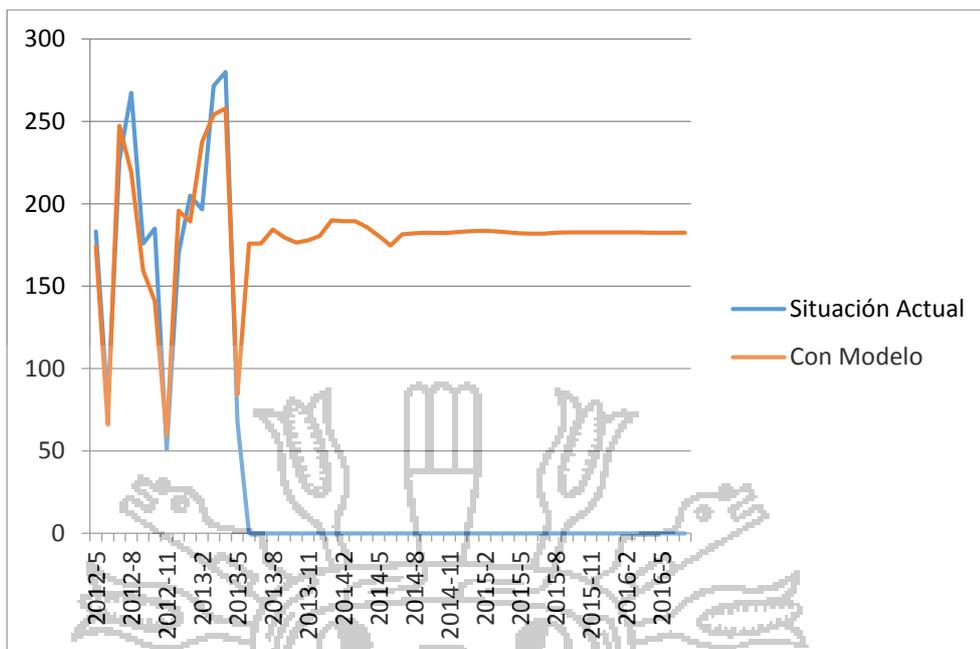


Figura 4.52: Predicción con Regresión Lineal ASO.

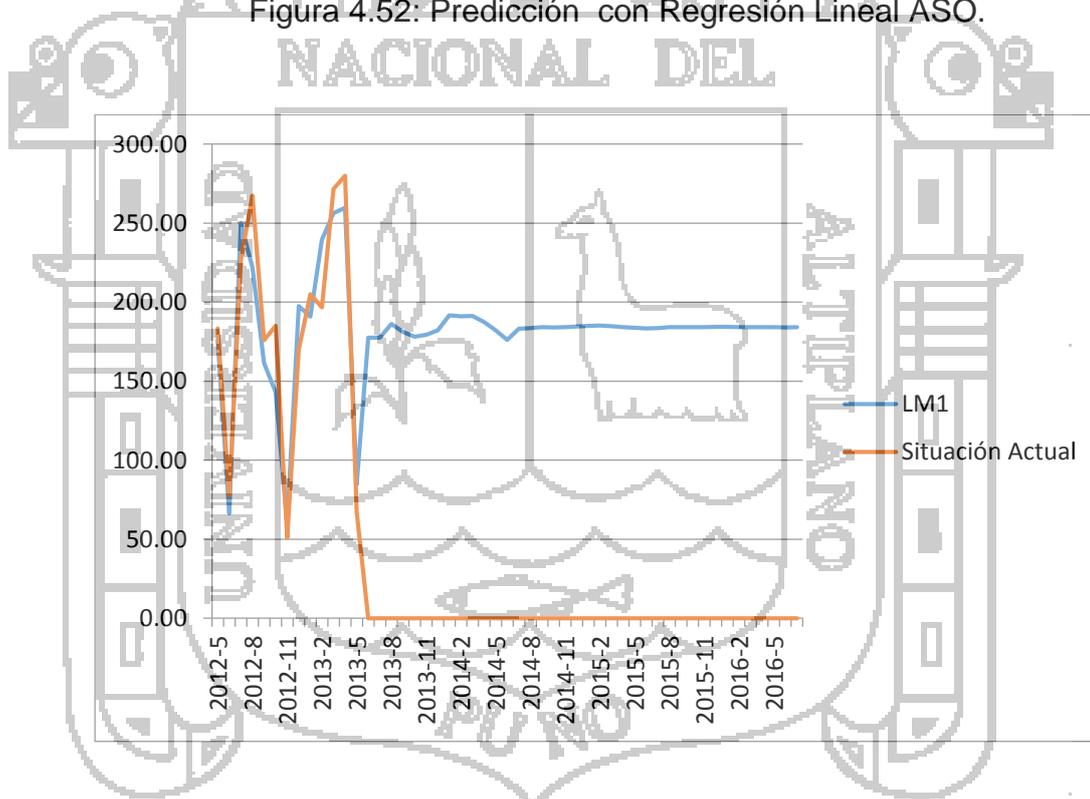


Figura 4.53: Predicción con Árboles de Decisión ASO.

Donde se puede observar que ASO está en función de NPO, PDP, PIPQ, PCA y PCR, con lo cual se obtiene la siguiente evaluación:

```

=== Evaluation on training set ===
=== Summary ===

Correlation coefficient           0.6128
Mean absolute error              3.2197
Root mean squared error          5.1245
Relative absolute error          70.2412 %
Root relative squared error      79.0235 %
Total Number of Instances       315

```

Figura 4.54: Evaluación con Regresión Lineal ASO.

```

=== Cross-validation ===
=== Summary ===

Correlation coefficient           0.5673
Mean absolute error              3.3751
Root mean squared error          5.3459
Relative absolute error          73.3292 %
Root relative squared error      82.1462 %
Total Number of Instances       315

```

Figura 4.55: Evaluación con Árboles de Decisión ASO

Esta prueba crea un modelo el cual obtiene un coeficiente de correlación de 0.6128, lo que significa que con regresión lineal clasifica correctamente el 61% de instancias.

d. Grado de Cumplimiento de las Operaciones Programadas (GCOP)

El modelo obtenido para este indicador es el siguiente:

```

Linear Regression Model

GCOP = -1 * PIQS + 100

Time taken to build model: 0 seconds

```

Figura 4.56: Modelo con Regresión Lineal GCOP en Unid.

```

M5 pruned model tree:

LM num: 1
GCOP = -1 * PIQS + 100

Number of Rules : 1

Time taken to build model: 0.03 seconds

```

Figura 4.57: Modelo con Árboles de Decisión GCOP en Unid.

A continuación se muestra el resultado de precisión de la predicción con el modelo hasta Julio del 2016 comparado con la situación actual hasta mayo del 2013 por lo que se concluye que del GCOP la Precisión de la predicción es precisa:

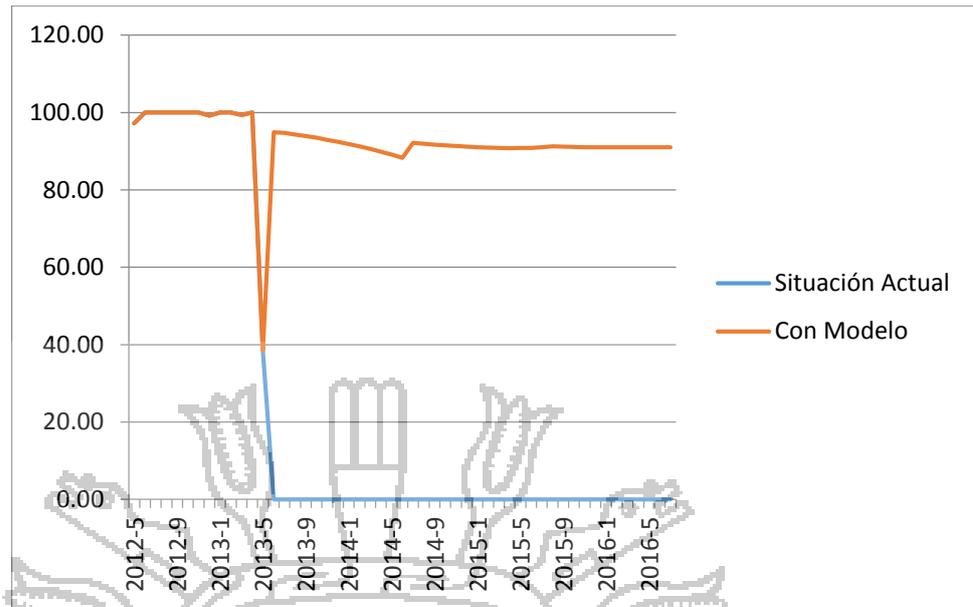


Figura 4.58: Predicción con Regresión Lineal GCOP.

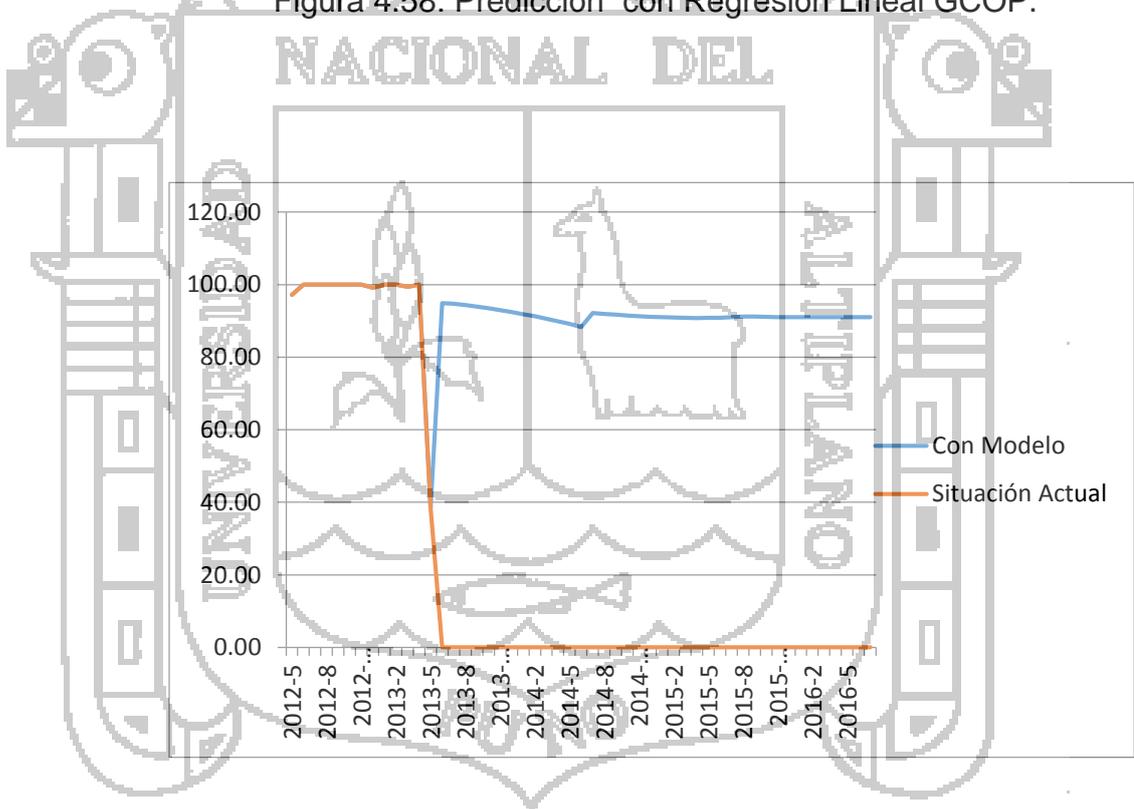


Figura 4.59: Predicción con Árboles de Decisión GCOP.

Donde se puede observar que GCOP está en función de PIQS, con lo cual se obtiene la siguiente evaluación:

```

=== Evaluation on training set ===

=== Summary ===

Correlation coefficient          1
Mean absolute error             0
Root mean squared error         0
Relative absolute error         0%
Root relative squared error     0%
Total Number of Instances      315

```

Figura 4.60: Evaluación con Regresión Lineal GCOP.

```

=== Cross-validation ===

=== Summary ===

Correlation coefficient          1
Mean absolute error             0
Root mean squared error         0
Relative absolute error         0%
Root relative squared error     0%
Total Number of Instances      315

```

Figura 4.61: Evaluación con Árboles de Decisión GCOP.

Esta prueba crea un modelo el cual obtiene un coeficiente de correlación de 1, lo que significa que clasifica correctamente con las dos técnicas el 100% de instancias.

e. Rendimiento Quirófano (RQ)

El modelo obtenido para este indicador es el siguiente:

```

Linear Regression Model
RQ = 0.2 * NIQ + 0
Time taken to build model: 0 seconds
    
```

Figura 4.62: Modelo con Regresión Lineal RQ en Unid.

```

M5 pruned model tree:
LM num: 1
RQ = 0.2 * NIQ + 0
Number of Rules : 1
Time taken to build model: 0.06 seconds
    
```

Figura 4.63: Modelo con Árboles de Decisión RQ en Unid.

A continuación se muestra el resultado de precisión de la predicción con modelo hasta julio el 2016 comparado con la situación actual hasta mayo del 2013 por lo que se concluye que la precisión de RQ es precisa:

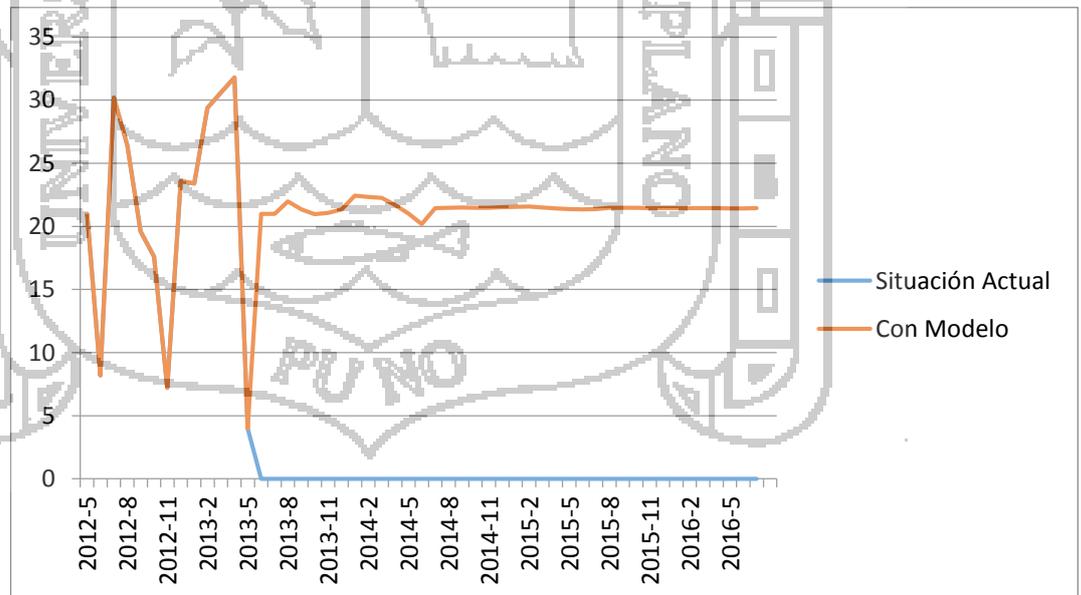


Figura 4.64: Predicción con Regresión Lineal RQ.

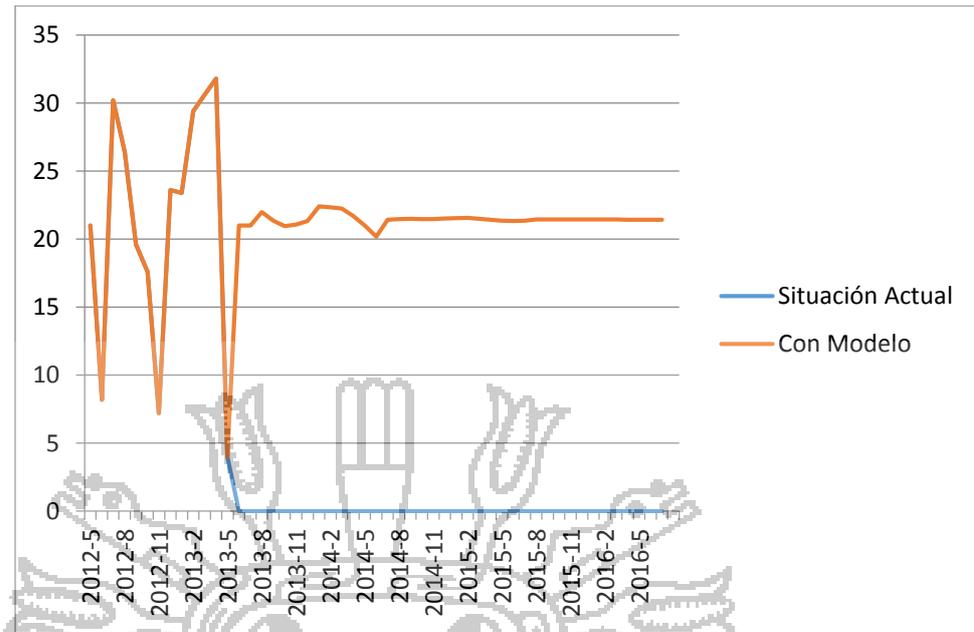


Figura 4.65: Predicción con Árboles de Decisión RQ.

Donde se puede observar que RQ está en función de NIQ, con lo cual se obtiene la siguiente evaluación:

```

=== Evaluation on training set ===
=== Summary ===

Correlation coefficient          1
Mean absolute error            0
Root mean squared error        0
Relative absolute error        0 %
Root relative squared error    0 %
Total Number of Instances     315
    
```

Figura 4.66: Evaluación con Regresión Lineal RQ.

```

=== Cross-validation ===
=== Summary ===

Correlation coefficient          1
Mean absolute error            0
Root mean squared error       0
Relative absolute error        0   %
Root relative squared error    0   %
Total Number of Instances     315

```

Figura 4.67: Evaluación con Árboles de Decisión RQ

Esta prueba crea un modelo el cual obtiene un coeficiente de correlación de 1, lo que significa que clasifica correctamente el 100% de instancias.

f. Porcentaje de Intervenciones Quirúrgicas Suspendidas (PIQS)

El modelo obtenido para este indicador es el siguiente:

```

Linear Regression Model

PIQS = -1 * GCOP + 100
Time taken to build model: 0.01 seconds

```

Figura 4.68: Modelo con Regresión Lineal PIQS en Unid.

```

M5 pruned model tree:

LM num: 1
PIQS = -1 * GCOP + 100
Number of Rules : 1
Time taken to build model: 0.03 seconds

```

Figura 4.69: Modelo con Árboles de Decisión PIQS en Unid.

A continuación se muestra el resultado de precisión de la predicción del modelo hasta Julio del 2016 comparado con la situación actual hasta mayo del 2013 por lo tanto se concluye que el PIPQ la precisión de la predicción es precisa:

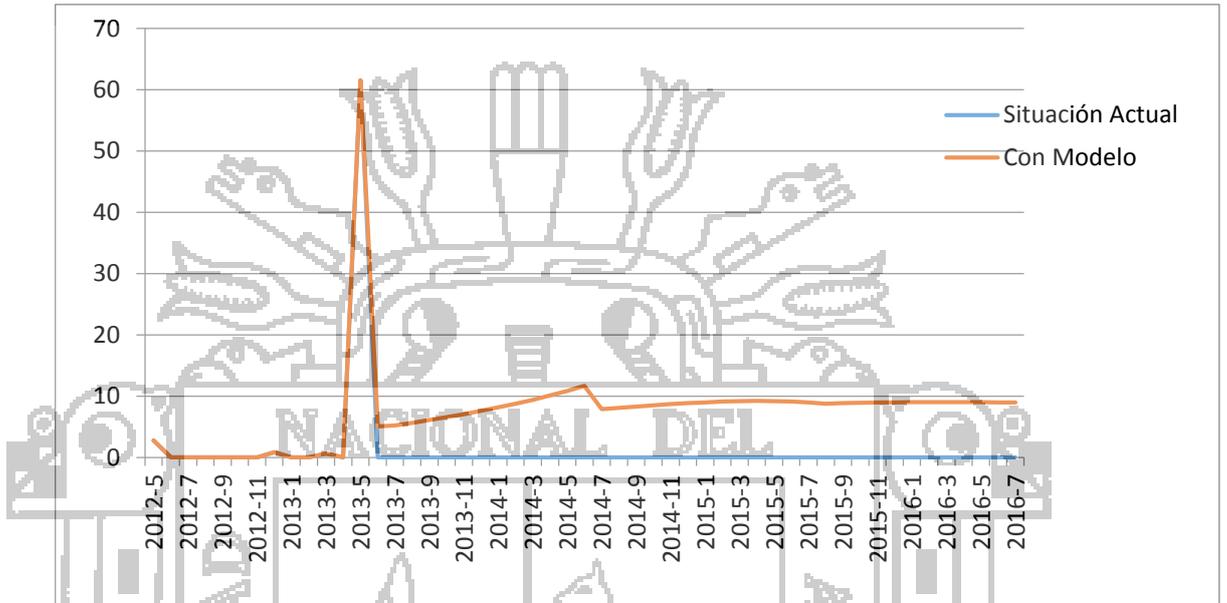


Figura 4.70: Predicción con Regresión Lineal PIQS

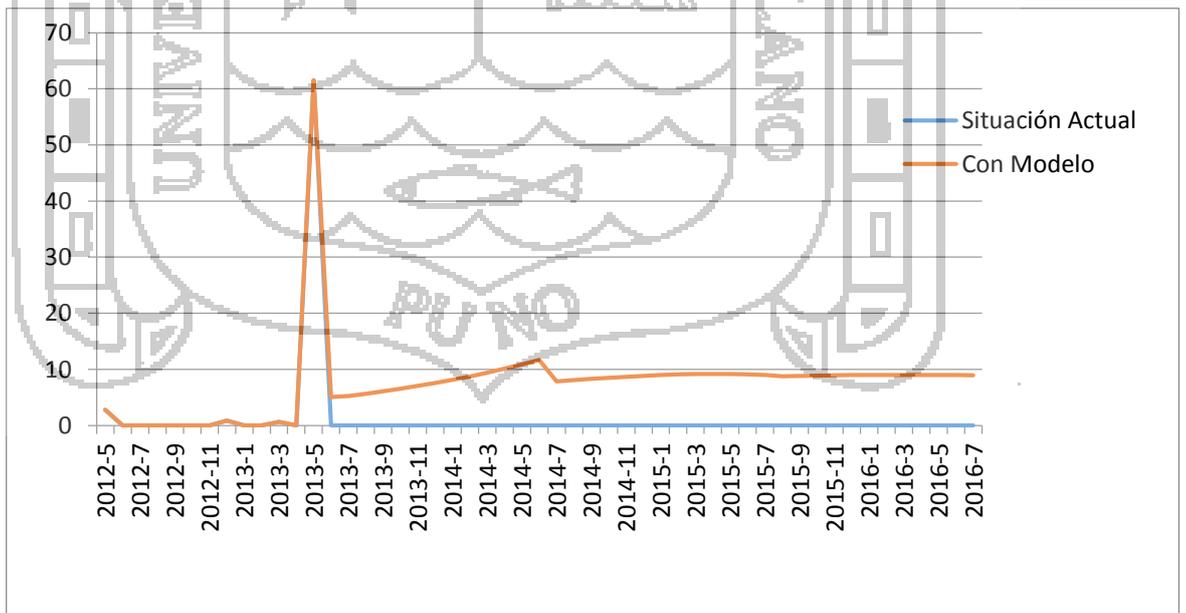


Figura 4.71: Predicción con Árboles de Decisión PIQS

Donde se puede observar que PIQS está en función de GCOP, con lo cual se obtiene la siguiente evaluación:

```

=== Evaluation on training set ===
=== Summary ===

Correlation coefficient          1
Mean absolute error            0
Root mean squared error        0
Relative absolute error        0    %
Root relative squared error    0    %
Total Number of Instances     315
  
```

Figura 4.72: Evaluación con Regresión Lineal PIQS.

```

=== Cross-validation ===
=== Summary ===

Correlation coefficient          1
Mean absolute error            0
Root mean squared error        0
Relative absolute error        0    %
Root relative squared error    0    %
Total Number of Instances     315
  
```

Figura 4.73: Evaluación con Árboles de Decisión PIQS.

Esta prueba crea un modelo el cual obtiene un coeficiente de correlación de 1, lo que significa que las dos técnicas clasifican correctamente el 100% de instancias.

g. Promedio de Días Preoperatorios (PDP)

El modelo obtenido para este indicador es el siguiente:

Linear Regression Model

$$\text{PDP} = 0.0283 * \text{ASO} - 0.0097 * \text{PIQS} - 0.1779 * \text{PCI} + 0.1768 * \text{PCA} - 0.3255 * \text{PCR} + 0.173 * \text{TM} + 0.8948$$

Time taken to build model: 0 seconds

Figura 4.74: Modelo con Regresión Lineal PDP en días.

M5 pruned model tree:

LM num: 1
 $\text{PDP} = 0.0101 * \text{NIQ} - 0.0126 * \text{NPO} + 0.0498 * \text{ASO} + 0.7283$

LM num: 2
 $\text{PDP} = 0.032 * \text{NIQ} - 0.1981 * \text{NPO} + 0.0076 * \text{ASO} + 2.4949$

Number of Rules : 2

Figura 4.75: Modelo con Árboles de Decisión PDP en días.

A continuación se muestra el resultado de precisión de la predicción con modelo hasta julio del 2016 comparado con la situación actual hasta mayo del 2013 en el que se concluye que no hay una precisión precisa del PDP ya que es directamente proporcional ASO PIQS, PCI y PCA, PCR:

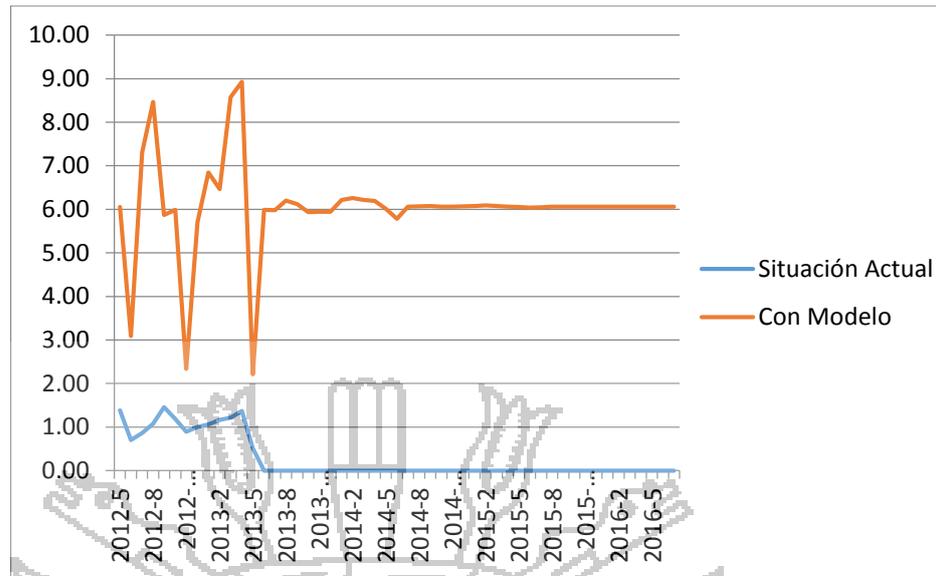


Figura 4.76: Predicción con Regresión Lineal PDP.

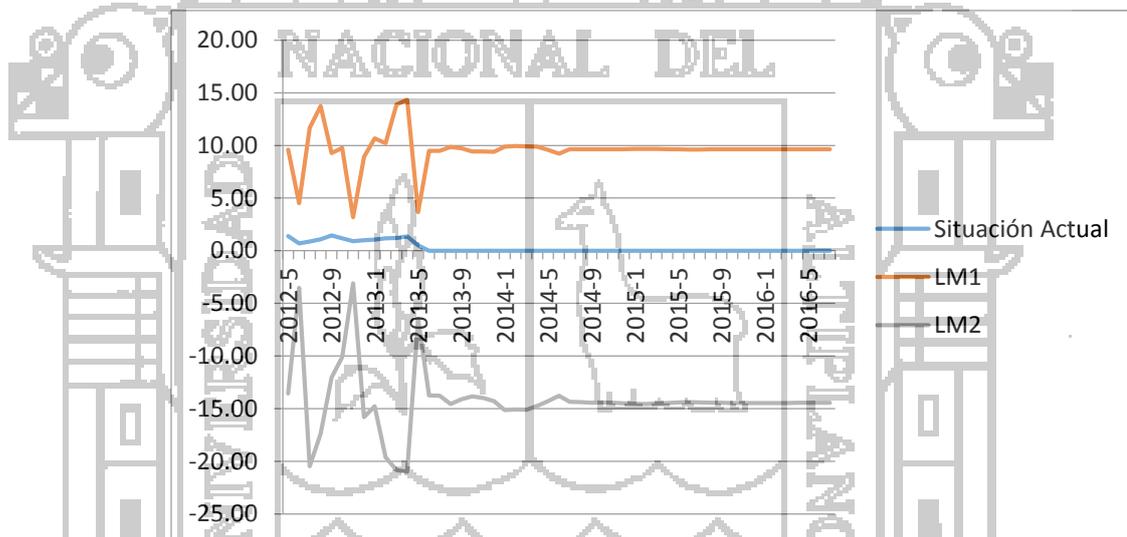


Figura 4.77: Predicción con Árboles de Decisión PDP.

Los valores negativos que se obtienen en la predicción de árboles de decisión de PDP con la regla 2 que es LM2 quiere decir que los datos no son suficientes para la aplicación de este algoritmo en la regla 2 por lo que no se considera en la evaluación del modelo dicha regla.

Se puede observar que PDP está en función de ASO, PIQS, PCI, PCA, PCR y TM, con lo cual se obtiene la siguiente evaluación:

```

=== Evaluation on training set ===
=== Summary ===

Correlation coefficient           0.257
Mean absolute error              0.9335
Root mean squared error         1.4796
Relative absolute error         97.2497 %
Root relative squared error     96.6416 %
Total Number of Instances       315

```

Figura 4.78: Evaluación con Regresión Lineal PDP.

```

=== Cross-validation ===
=== Summary ===

Correlation coefficient           0.0539
Mean absolute error              0.974
Root mean squared error         1.5476
Relative absolute error         101.0304 %
Root relative squared error     100.7805 %
Total Number of Instances       315

```

Figura 4.79: Evaluación con Árboles de Decisión PDP.

Esta prueba crea un modelo el cual obtiene un coeficiente de correlación de 0.257, lo que significa que con regresión lineal clasifica correctamente el 25.7% de instancias.

h. Porcentaje de Infecciones Post Quirúrgicas (PIPQ)

El modelo obtenido para este indicador es el siguiente:

Linear Regression Model

```

PIPQ = -53.4782 * Anio - 1.5744 * NIQ + 0.5663 * ASO - 0.1959
      * PIQS - 1.1749 * PCI + 1.0802 * PCA + 107660.1428
Time taken to build model: 0 seconds

```

Figura 4.80: Modelo con Regresión Lineal PIPQ en Unid.

```

M5 pruned model tree:
LM num: 1
PIPQ = -4.1412 * Anio + 6.7216 * Mes + 0.8829 * Dia - 0.3147 * NIQ
      + 0.1117 * ASO + 8308.0566
LM num: 2
PIPQ = -4.1412 * Anio + 16.2367 * Mes + 2.9414 * Dia - 0.3147 *
      NIQ + 0.1117 * ASO + 8213.1586
LM num: 3
PIPQ = -4.1412 * Anio + 16.2367 * Mes + 3.0873 * Dia - 0.3147 *
      NIQ + 0.1117 * ASO + 8213.4936
LM num: 4
PIPQ = -4.1412 * Anio + 15.6506 * Mes + 1.9045 * Dia - 0.3147 *
      NIQ + 0.1117 * ASO + 8260.255
LM num: 5
PIPQ = -4.1412 * Anio + 1.681 * Mes + 0.1259 * Dia - 0.3147 *
      NIQ + 0.1117 * ASO + 8407.8131
LM num: 6
PIPQ = -4.1412 * Anio - 2.0063 * Mes + 0.5296 * Dia - 0.3147 *
      NIQ + 0.1117 * ASO + 8427.3189
LM num: 7
PIPQ = -4.1412 * Anio - 3.4584 * Mes + 2.3112 * Dia - 0.3147 *
      NIQ + 0.1117 * ASO + 8391.0642
LM num: 8
PIPQ = -4.1412 * Anio + 1.5042 * Dia - 0.944 * NIQ + 0.335 * ASO + 8359.6183
LM num: 9
PIPQ = -4.1412 * Anio - 0.272 * Dia - 0.944 * NIQ + 0.335 * ASO + 8362.4186
LM num: 10
PIPQ = -5.7804 * Anio - 0.3164 * Mes + 0.2898 * NIQ - 0.0855 * ASO + 11637.8464
LM num: 11
PIPQ = -5.7804 * Anio - 1.8601 * Mes + 0.1109 * NIQ - 0.0327 * ASO + 11645.5979
Number of Rules : 11
Time taken to build model: 0.06 seconds

```

Figura 4.81: Modelo con Árboles de Decisión PIPQ en Unid.

A continuación se muestra el resultado de la precisión de predicción con modelo hasta julio el 2016 comparado con la situación actual hasta mayo del 2013 en el que se deduce que el PIPQ depende de NIQ:

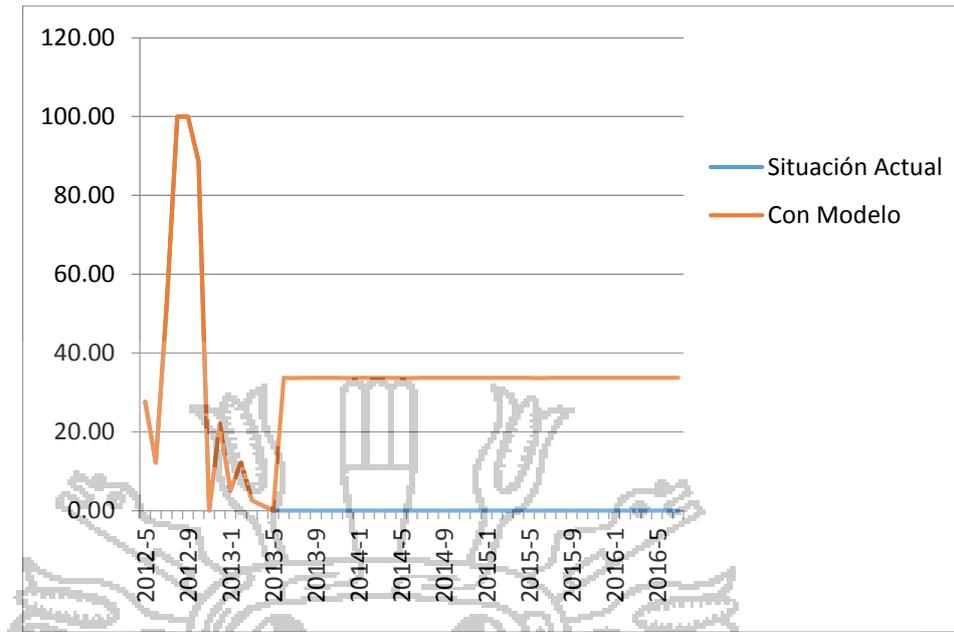


Figura 4.82: Predicción con Regresión Lineal PIPQ.

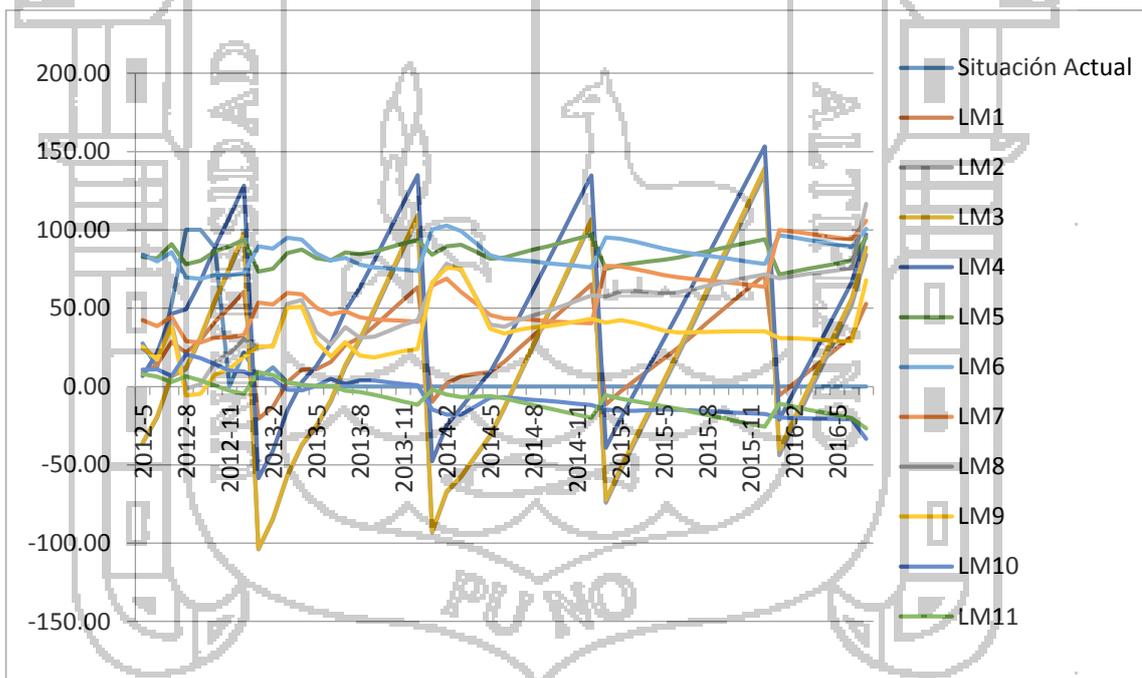


Figura 4.83: Predicción con Árboles de Decisión PIPQ.

Los valores negativos que se obtienen en la predicción de árboles de decisión de PIPQ con las 11 reglas quiere decir que los datos no son

suficientes para la aplicación de este algoritmo por lo que no se considera en la evaluación del modelo de dichas reglas los valores negativos.

Se puede observar que PIPQ está en función de NIQ, ASO, PIQS, PCI y PCA, con lo cual se obtiene la siguiente evaluación:

```

=== Evaluation on training set ===
=== Summary ===
Correlation coefficient      0.6132
Mean absolute error        28.9089
Root mean squared error    35.6674
Relative absolute error    68.642 %
Root relative squared error 78.9952 %
Total Number of Instances  315
  
```

Figura 4.84: Evaluación con Regresión Lineal PIPQ.

```

=== Cross-validation ===
=== Summary ===
Correlation coefficient      0.9035
Mean absolute error        14.1407
Root mean squared error    19.8384
Relative absolute error    33.3447 %
Root relative squared error 43.6244 %
Total Number of Instances  315
  
```

Figura 4.85: Evaluación con Árboles de Decisión PIPQ.

Esta prueba crea un modelo el cual obtiene un coeficiente de correlación de 0.6132, lo que significa que con regresión lineal clasifica correctamente el 61.32% de instancias pero con árboles de decisión con un coeficiente de correlación de 0.9035 lo que quiere decir que clasifica con un 90%.

i. Porcentaje de Complicaciones Intraoperatorias (PCI)

El modelo obtenido para este indicador es el siguiente:

```

Linear Regression Model
PCI = -0.0481 * PDP + 0.9998 * PCA - 1.187 * PCR +
      0.2522 * TM + 0.0704
Time taken to build model: 0 seconds
  
```

Figura 4.86: Modelo con Regresión Lineal PCI en Unid.

```

M5 pruned model tree:
LM num: 1
PCI = + 0.3704
Number of Rules : 1
Time taken to build model: 0.02 seconds
  
```

Figura 4.87: Modelo con Árboles de Decisión PCI en Unid.

A continuación se muestra el resultado de la precisión de la predicción con modelo hasta el 2016 comparado con la situación actual hasta mayo

del 2013 por lo que se concluye que la precisión de la predicción del PCI es inversamente proporcional al PCA y PCR:

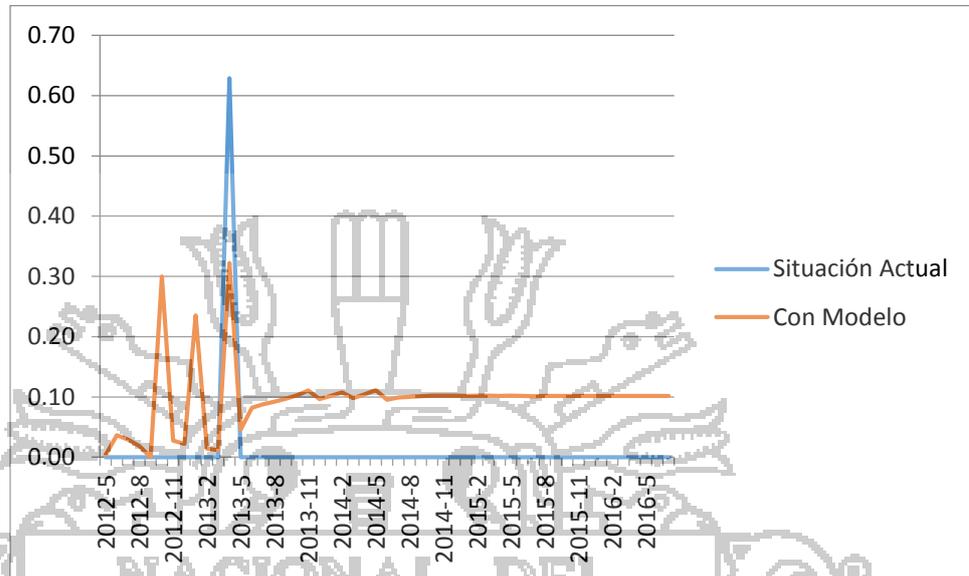


Figura 4.88: Predicción con Regresión Lineal PCI.

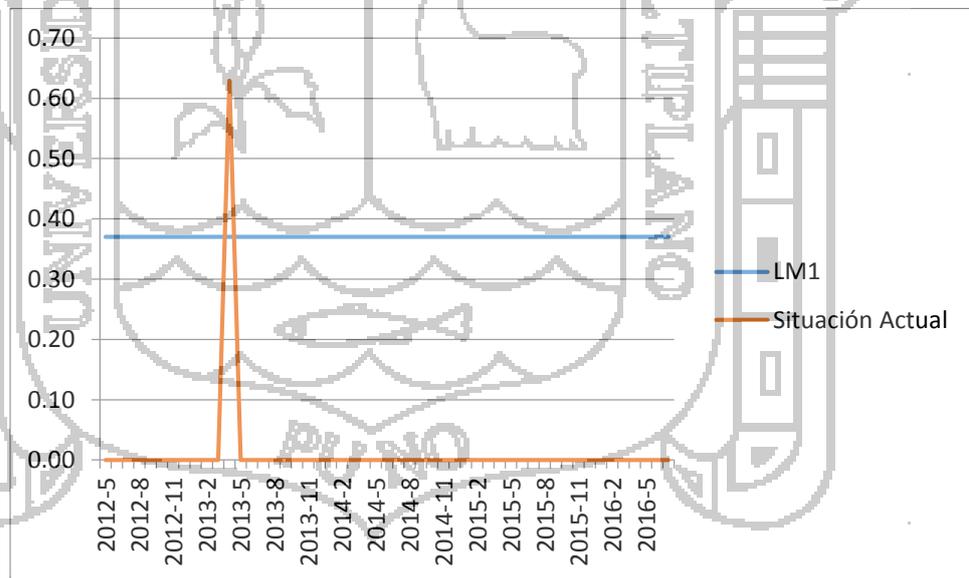


Figura 4.89: Predicción con Árboles de Decisión PCI.

Donde se puede observar que PCI está en función de PDP, PCA, PCR y TM, con lo cual se obtiene la siguiente evaluación:

```

=== Evaluation on training set ===

=== Summary ===

Correlation coefficient           0.9898
Mean absolute error              0.1365
Root mean squared error          0.8119
Relative absolute error          18.5417 %
Root relative squared error      14.2428 %
Total Number of Instances       315

```

Figura 4.90: Evaluación con Regresión Lineal PCI.

```

=== Cross-validation ===
=== Summary ===

Correlation coefficient           -0.0117
Mean absolute error              1.0738
Root mean squared error          6.0217
Relative absolute error          145.5721 %
Root relative squared error      105.3028 %
Total Number of Instances       315

```

Figura 4.91: Evaluación con Árboles de Decisión PCI.

Esta prueba crea un modelo el cual obtiene un coeficiente de correlación de 0.9898, lo que significa que clasifica con regresión lineal correctamente el 98.98% de instancias.

j. Porcentaje de Complicaciones Anestésicas (PCA)

El modelo obtenido para este indicador es el siguiente:

Linear Regression Model

$$PCA = 0.047 * PDP + 0.9798 * PCI + 1.1834 * PCR - 0.2474 * TM - 0.0622$$

Time taken to build model: 0 seconds

Figura 4.92: Modelo con Regresión Lineal PCA en Unid.

M5 pruned model tree:

LM num: 1

$$PCA = 0.2099 * Mes - 0.8225$$

Number of Rules : 1

Time taken to build model: 0.02 seconds

Figura 4.93: Modelo con Árboles de Decisión PCA en Unid.

A continuación se muestra el resultado de la precisión de predicción con modelo hasta julio del 2016 comparado con la situación actual hasta mayo del 2013 en el que se concluye que el PCA es directamente proporcional al PCI:

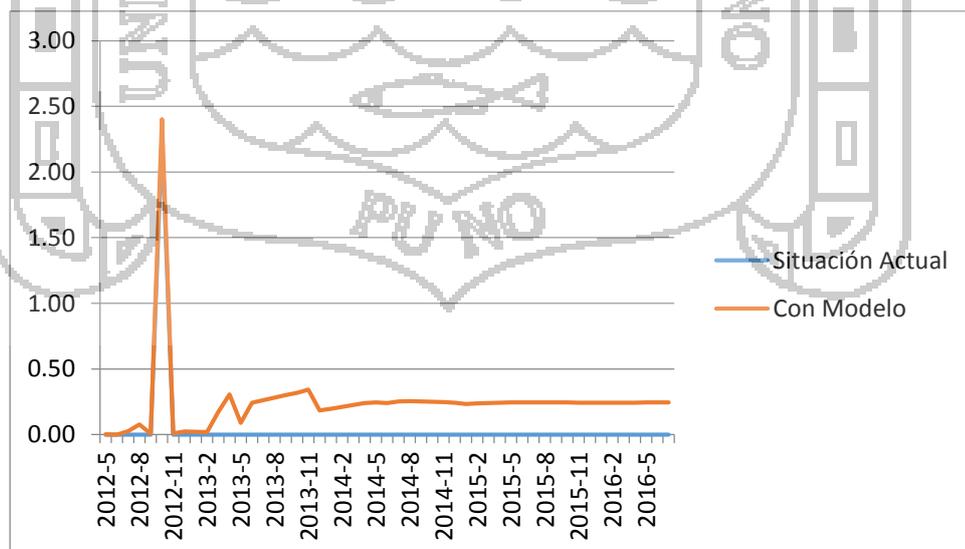


Figura 4.94: Predicción con Regresión Lineal PCA.

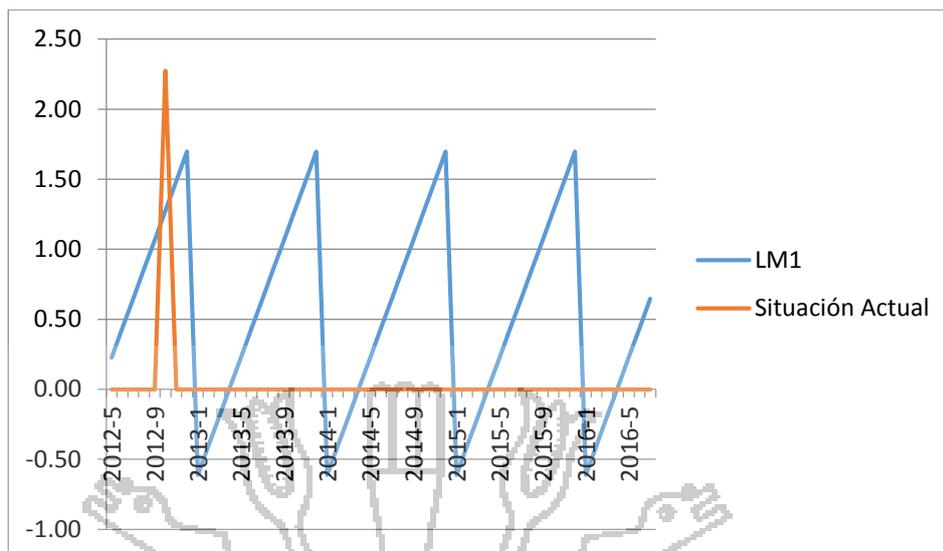


Figura 4.95: Predicción con Árboles de Decisión PCA.

Los valores negativos que se obtienen en la predicción de árboles de decisión de PCA con la regla 1 de LM1 quiere decir que los datos no son suficientes para la aplicación de este algoritmo para dicha regla por lo que no se considera en la evaluación del modelo los datos negativos de dicha regla sólo los valores positivos.

Donde se puede observar que PCA está en función de PDP, PCI, PCR y TM, con lo cual se obtiene la siguiente evaluación:

```

=== Evaluation on training set ===
=== Summary ===
Correlation coefficient          0.9911
Mean absolute error            0.1365
Root mean squared error        0.8037
Relative absolute error        14.1602 %
Root relative squared error    13.3384 %
Total Number of Instances      315
    
```

Figura 4.96: Evaluación con Regresión Lineal PCA.

```

=== Cross-validation ===
=== Summary ===

Correlation coefficient           -0.0174
Mean absolute error              0.9974
Root mean squared error         6.0792
Relative absolute error         103.075 %
Root relative squared error     100.5493 %
Total Number of Instances      315

```

Figura 4.97: Evaluación con Árboles de Decisión PCA.

Esta prueba crea un modelo el cual obtiene un coeficiente de correlación de 0.9911, lo que significa que clasifica correctamente el 99.11% de instancias.

k. Porcentaje de Complicaciones en Recuperación (PCR)

El modelo obtenido para este indicador es el siguiente:

```

Linear Regression Model

PCR = -0.1542 * PDP + 0.0043 * PIPQ +
      -0.5799 * TM + 0.0236

Time taken to build model: 0 seconds

```

Figura 4.98: Modelo con Regresión Lineal PCR en porcentaje.

```

M5 pruned model tree:
LM num: 1
PCR =0.0074 * PCA + 0.006
LM num: 2
PCR =-0.3256 * Mes + 0.5789 * PCA + 3.7788
LM num: 3
PCR =-0.2428 * Mes + 0.0524 * PCA + 2.8219
Number of Rules : 3
Time taken to build model: 0.02 seconds
    
```

Figura 4.99: Modelo con Árboles de Decisión PCR en porcentaje.

A continuación se muestra el resultado de la precisión de predicción hasta julio del 2016 comparado con el modelo hasta mayo del 2013 en el que se concluye que el PCR depende de TM si la TM aumenta también aumenta el PCR:

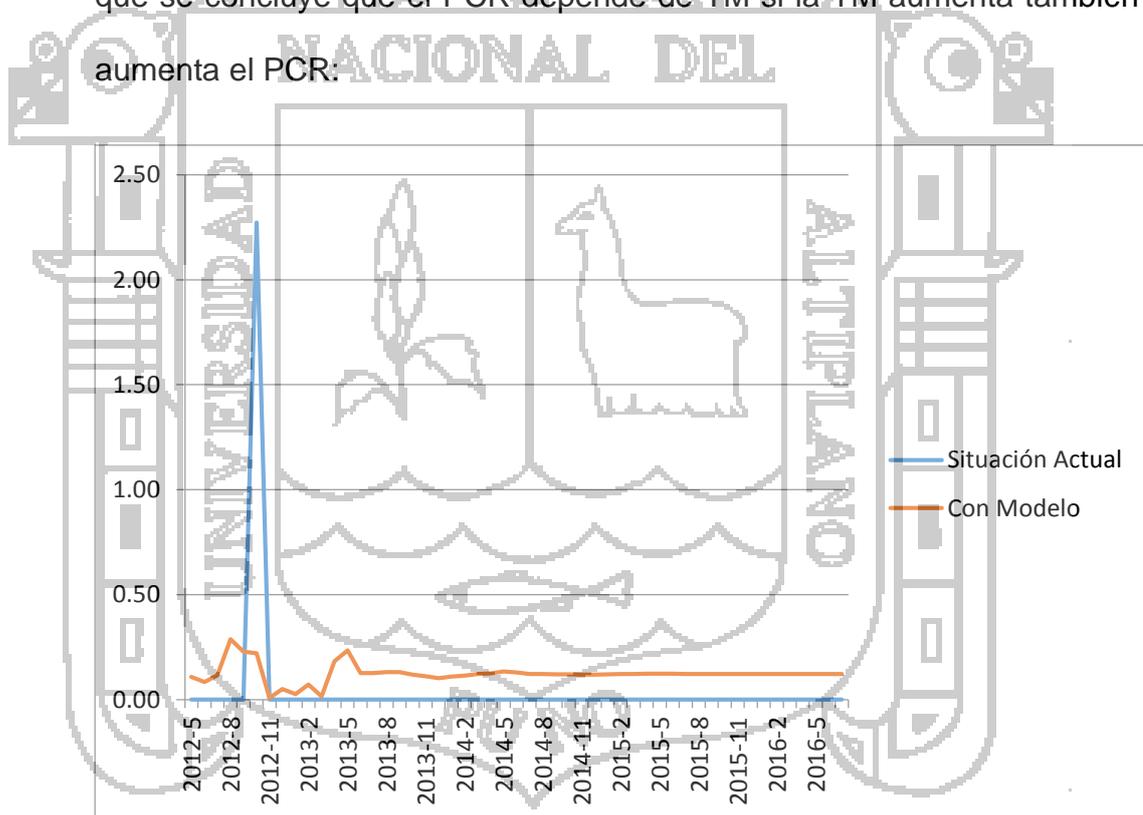


Figura 4.100: Predicción con Regresión Lineal PCR.

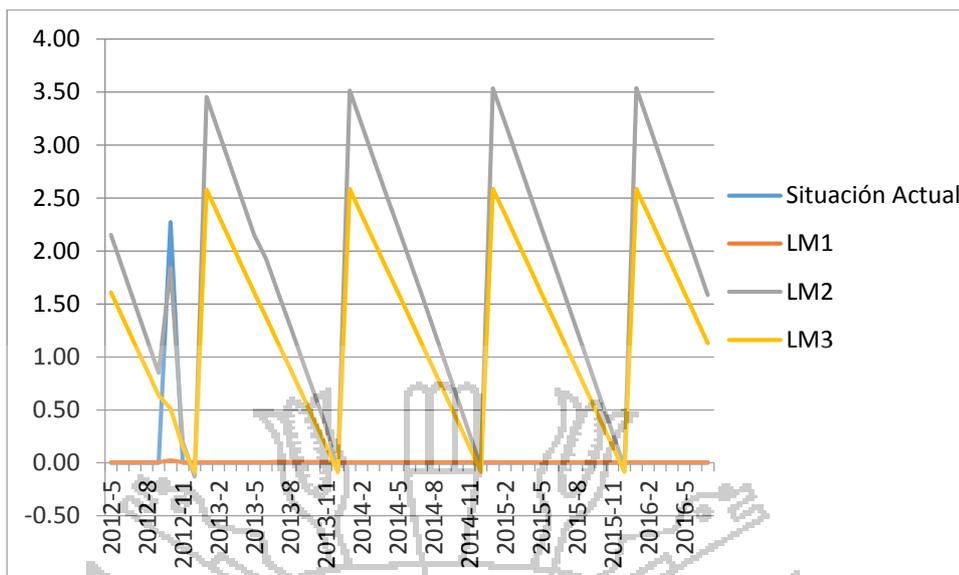


Figura 4.101: Predicción con Árboles de Decisión PCR.

Los valores negativos que se obtienen en la predicción de árboles de decisión de PCR de la regla 3 de LM3 quiere decir que los datos no son suficientes para la aplicación de este algoritmo para dicha regla por lo que no se considera en la evaluación del modelo los datos negativos de dicha regla sólo los valores positivos.

Donde se puede observar que PCR también está en función de PDP, PIPQ, y TM, con lo cual se obtiene la siguiente evaluación:

```

=== Evaluation on training set ===
=== Summary ===
Correlation coefficient           0.6601
Mean absolute error              0.3937
Root mean squared error          1.6403
Relative absolute error          117.0103 %
Root relative squared error       75.1147 %
Total Number of Instances        315
    
```

Figura 4.102: Evaluación con Regresión Lineal PCR.

```

=== Cross-validation ===
=== Summary ===

Correlation coefficient           -0.005
Mean absolute error              0.5406
Root mean squared error          6.0473
Relative absolute error          159.4889 %
Root relative squared error      275.1908 %
Total Number of Instances        315

```

Figura 4.103: Evaluación con Árboles de Decisión PCR.

Esta prueba crea un modelo el cual obtiene un coeficiente de correlación de 0.6601, lo que significa que con regresión lineal clasifica correctamente el 66.01% de instancias.

I. Tasa de Mortalidad (TM)

El modelo obtenido para este indicador es el siguiente:

```

Linear Regression Model
TM = 0.3952 * Anio + 0.2099 * PDP + 0.7411 * PCR -795.3897
Time taken to build model: 0 seconds

```

Figura 4.104: Modelo con Regresión Lineal TM en Porcentaje.

```

M5 pruned model tree:
LM num: 1
TM = 0.1784 * PDP + 0.07
Number of Rules : 1
Time taken to build model: 0.02 seconds

```

Figura 4.105: Modelo con Árboles de Decisión TM en Porcentaje.

A continuación se muestra el resultado de la precisión de predicción hasta julio del 2016 con el modelo hasta mayo 2013 por lo que se concluye que TM depende de PDP y PCR si estos son altos la TM aumenta:

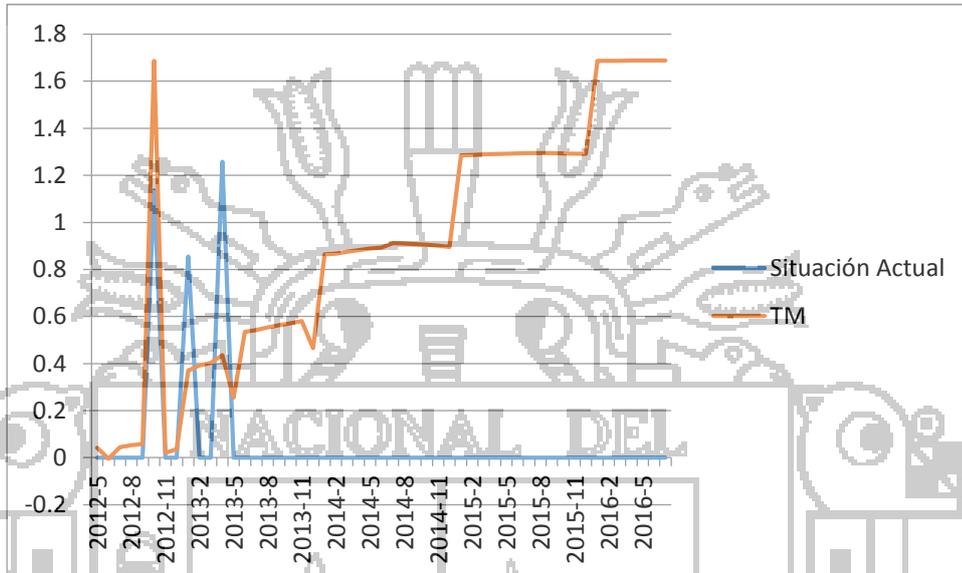


Figura 4.106: Predicción con Regresión Lineal TM.

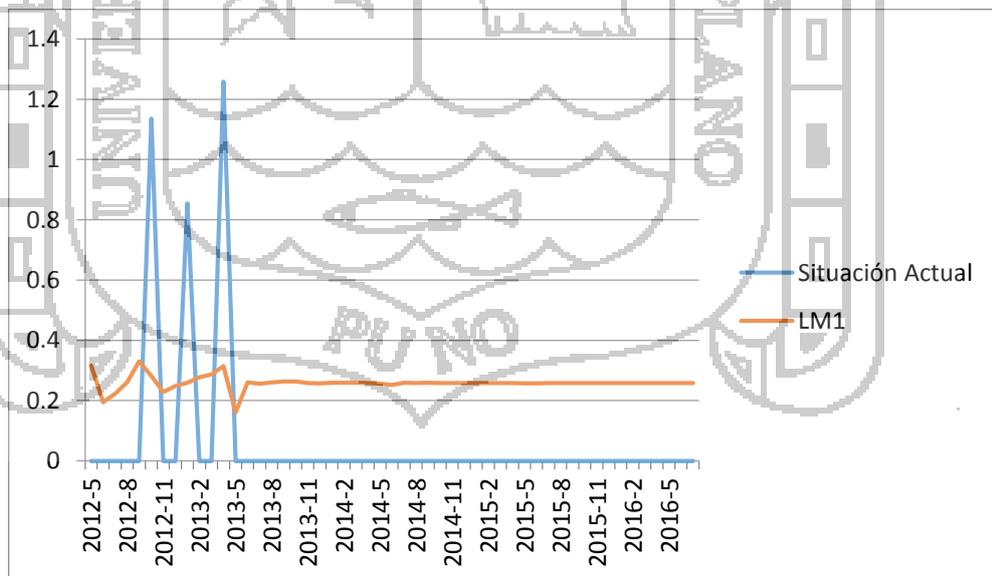


Figura 4.107: Predicción con Árboles de Decisión TM.

Donde se puede observar que TM está en función de PDP y PCR, con lo cual se obtiene la siguiente evaluación:

```

=== Evaluation on training set ===
=== Summary ===

Correlation coefficient           0.6635
Mean absolute error              0.4871
Root mean squared error         1.8489
Relative absolute error         94.1905 %
Root relative squared error     74.8188 %
Total Number of Instances      315
  
```

Figura 4.108: Evaluación con Regresión Lineal TM.

```

=== Cross-validation ===
=== Summary ===

Correlation coefficient          -0.0512
Mean absolute error             0.5231
Root mean squared error         2.5048
Relative absolute error        100.8715 %
Root relative squared error    101.1156 %
Total Number of Instances      315
  
```

Figura 4.109: Evaluación con Árboles de Decisión TM.

Esta prueba crea un modelo el cual obtiene un coeficiente de correlación de 0.6635, lo que significa que clasifica correctamente el 66.35% de instancias.

4.3.8. Evaluación Técnica del Modelo

La siguiente tabla presenta un resumen de los resultados obtenidos en la construcción de los modelos de regresión lineal para los indicadores hospitalarios.

Tabla 4.13: Evaluación Técnica del Modelo.

| Indicador | Coefficiente correlación | Error en la media absoluta | Error absoluto relativo |
|-----------------|--------------------------|----------------------------|-------------------------|
| NIQ | 1 | 0 | 0% |
| NPO | 0.9786 | 0.1313 | 6.88% |
| ASO | 0.6128 | 3.2197 | 70.24% |
| GCOP | 1 | 0 | 0 |
| RQ | 1 | 0 | 0 |
| PIQS | 1 | 0 | 0 |
| PDP | 0.257 | 0.9335 | 97.2497 |
| PIP | 0.6132 | 28.9089 | 68.642 |
| PCI | 0.9898 | 0.1365 | 18.5417 |
| PCA | 0.9911 | 0.1365 | 14.1602 |
| PCR | 0.6601 | 0.3937 | 117.0103 |
| TM | 0.6635 | 0.4871 | 94.1905 |
| Promedio | 0.8138 | 2.6823 | 34.2138 |

Fuente: Elaboración Propia

Para la evaluación del modelo en general se estableció un promedio entre todos los indicadores, teniendo como resultado un coeficiente de correlación de 0.8138 lo que significa que el 81.38% de instancias fueron clasificadas correctamente para este conjunto de pruebas.

4.4. Evaluación del Modelo

En esta etapa se realizó una evaluación técnica final en función a los objetivos del departamento de cirugía y una revisión del proceso de modelado.

4.4.1. Evaluación de los Resultados

Al aplicar la técnica de regresión lineal sobre el conjunto de datos se han encontrado observaciones sobre un indicador hospitalario, cuyo modelo no es recomendable para la predicción.

El conjunto de datos de entrenamiento permitió al algoritmo mejorar la tasa de aciertos para los indicadores.

En el siguiente cuadro se detalla el cumplimiento del criterio de éxito establecido para el proyecto.

Tabla 4.14: Evaluación de los Resultados

| Objetivos del Negocio | Objetivo de minería de datos | Criterio de éxito |
|----------------------------------|---------------------------------------|---|
| Mejora en la toma de decisiones. | Control de indicadores hospitalarios. | 81.38%, mayor al grado de precisión mínimo. |

Fuente: Elaboración Propia

Sin embargo es importante destacar que este modelo es notablemente acertero si se dispusiera de mayor cantidad de datos sería mucho mejor, es decir más registros de informes operatorios de años anteriores con los datos completos.

El modelo aprobado resultante es un conjunto de ecuaciones de regresión lineal correspondientes a los indicadores hospitalarios, construidas a partir de las variables definidas en la selección y limpieza de datos. El coeficiente de correlación para el modelo fue de 81.38%.

4.4.2. Revisión del Proceso

Se cumplieron todos los pasos que involucran el proceso de minería de datos para la construcción del modelo para cada uno de los indicadores hospitalarios definidos como meta del proyecto.

Sin embargo en la fase de recolección de datos se tuvo que construir un software para registrar los datos en el sistema, por lo que se invirtió mucho tiempo en esta etapa.

4.4.3. Prueba de Hipótesis

Para la prueba de hipótesis se utilizan las formulas descritas en el diseño de la prueba de hipótesis (punto 3.9), como sigue:

a. Predicción:

Para realizar la prueba de la predicción se trabajan con los siguientes datos, obteniendo el porcentaje de precisión por cada indicador:

Los datos para la predicción fueron hallados de la siguiente manera ver tabla 4.15:

Tabla 4.15 Datos de Precisión de la Predicción.

| Indicadores | % Datos Sin Modelo | % Datos Con Modelo |
|--|---------------------------|---------------------------|
| Número de Intervenciones Quirúrgicas | 0% | 100 |
| Número de Pacientes Operados | 0% | 94.92063492 |
| Aprovechamiento Sala de Operaciones | 0% | 15.87301587 |
| Grado de Cumplimiento de las Operaciones Programadas | 0% | 100 |
| Rendimiento de Quirófano | 0% | 100 |
| Porcentaje de Intervenciones Quirúrgicas Suspendidas | 0% | 100 |
| Promedio de Días preoperatorios | 0% | 33.65079365 |
| Porcentaje de Infecciones Post Quirúrgicas | 0% | 100 |
| Porcentaje de Complicaciones Intraoperatorias | 0% | 98.41269841 |
| Porcentaje de Complicaciones Anestésicas | 0% | 98.0952381 |
| Porcentaje de Complicaciones en Recuperación | 0% | 94.6031746 |
| Tasa de mortalidad | 0% | 71.74603175 |
| TOTAL | 0% | 83.94179894 |

Fuente: Elaboración Propia.

Del cuadro se tiene que la cantidad de datos con precisión de la predicción es 83.94%.

Aplicando la fórmula de Datos de la Precisión de la Predicción se tiene que con el modelo se mejora en un 100%.

b. Porcentaje de precisión:

Para realizar la prueba de la precisión se trabajan con los siguientes datos, obteniendo el porcentaje de precisión por cada indicador hospitalario.

Los datos para la precisión fueron hallados de la siguiente manera ver tabla 4.16:

Tabla 4.16: Resumen del Porcentaje Precisión

| Indicadores Hospitalarios | Precisión Con Modelo | Precisión Sin Modelo | Incorrecto | % de Precisión |
|--|-----------------------------|-----------------------------|-------------------|-----------------------|
| Número de Intervenciones quirúrgicas | 315 | 315 | 0 | 100 |
| Número de Pacientes Operados | 315 | 299 | 16 | 94.92063492 |
| Aprovechamiento Sala de Operaciones | 315 | 50 | 265 | 15.87301587 |
| Grado de Cumplimiento de las Operaciones Programadas | 315 | 315 | 0 | 100 |
| Rendimiento de Quirófano | 315 | 315 | 0 | 100 |
| Porcentaje de Intervenciones Quirúrgicas Suspendidas | 315 | 315 | 0 | 100 |
| Promedio de Días Preoperatorios | 315 | 106 | 209 | 33.65079365 |
| Porcentaje de Infecciones Post Quirúrgicas | 315 | 315 | 0 | 100 |
| Porcentaje de Complicaciones Intraoperatorias | 315 | 310 | 5 | 98.41269841 |
| Porcentaje de Complicaciones Anestésicas | 315 | 309 | 6 | 98.0952381 |
| Porcentaje de Complicaciones en Recuperación | 315 | 298 | 17 | 94.6031746 |
| Tasa de Mortalidad | 315 | 226 | 89 | 71.74603175 |
| Total | 3780 | 3173 | 607 | 83.94179894 |

Fuente: Elaboración Propia

De la tabla anterior se tiene que el porcentaje de precisión cumple con un 83.94% de precisión.

c. Datos históricos:

Los datos históricos fueron hallados de la siguiente manera:

Tabla 4.17: Datos Históricos.

| Indicadores | % Datos Históricos Sin Modelo | % Datos Históricos Con Modelo |
|--|--------------------------------------|--------------------------------------|
| Número de Intervenciones Quirúrgicas | SI | SI |
| Número de Pacientes Operados | SI | SI |
| Aprovechamiento Sala de Operaciones | SI | SI |
| Grado de Cumplimiento de las Operaciones Programadas | NO | SI |
| Rendimiento de Quirófano | SI | SI |
| Porcentaje de Intervenciones Quirúrgicas Suspendidas | SI | SI |
| Promedio de Días preoperatorios | NO | SI |
| Porcentaje de Infecciones Post Quirúrgicas | NO | SI |
| Porcentaje de Complicaciones Intraoperatorias | NO | SI |
| Porcentaje de Complicaciones Anestésicas | NO | SI |
| Porcentaje de Complicaciones en Recuperación | NO | SI |
| Tasa de mortalidad | NO | SI |
| TOTAL | 6 | 12 |

Fuente: Elaboración Propia.

Del cuadro se tiene que la cantidad de datos históricos que se tiene actualmente sin el modelo son de 6 indicadores hospitalarios ya que la información que se almacena es en formatos preoperatorios como se explicó en el punto 1.1. y con el modelo se llega a obtener información de los 12 indicadores hospitalarios ya que la información se almacena en un sistema de información automatizado.

Aplicando la fórmula de Datos Históricos se tiene que con el modelo se mejora en un 100%.

d. Mejora en la Toma de Decisiones

De los datos obtenidos se tiene:

$$\text{Mejora en la toma de decisiones} = \frac{83.94 + 100}{2} = 91.97\%$$

Figura 4.110: Fórmula Mejora en la Toma de Decisiones.

Por lo tanto el modelo de minería de datos mejora en la toma de decisiones en un porcentaje del 91.97%, con un grado excelente ya que está en el rango de 76-100%.



CONCLUSIONES

PRIMERA

Se desarrolló el Modelo Predictivo Basado en Minería de Datos para la Mejora en la Toma de Decisiones del Departamento de Cirugía del Hospital Regional Manuel Nuñez Butrón, gracias al promedio del Coeficiente de Correlación de los indicadores hospitalarios de 81.34% que se obtuvo la evaluación del modelo es correcta mejorando así en la toma de decisiones en un porcentaje de 91.97%, Ver el punto 4.4.3

SEGUNDA

Se logró comprender la situación actual del departamento de cirugía del Hospital Regional Manuel Núñez Butrón identificando los objetivos del departamento de cirugía, definiendo el problema de la minería de datos, identificando como caso de éxito el índice de precisión para la mejora en la toma de decisiones y se definió el plan de trabajo de 15 semanas. Ver punto 4.1.

TERCERA

Se comprendió los datos del departamento de cirugía del Hospital regional Manuel Núñez Butrón familiarizándose y recolectando los mismos desarrollando un software para el modelado y su correspondiente almacenamiento; así como la arquitectura de funcionamiento, 4 casos de uso, 4 diagramas de secuencia, 4 diagramas de actividades y 1 diagrama de clases,

se validó la calidad de los datos y finalmente se definió la hipótesis en el que el índice de precisión sea mayor o igual al 75% para el criterio de éxito en la mejora de la toma de decisiones. Ver punto 4.2.

CUARTA

Se procesaron los datos para el modelo logrando identificar 12 indicadores hospitalarios para la evaluación del centro quirúrgico, de esta manera se seleccionaron los datos necesarios de acuerdo a la fórmula de cada indicador, en el que se realizó el filtrado de datos en MYSQL finalmente se procesaron 1500 registros obtenidos desde mayo del 2012 hasta mayo del 2013. Ver punto 4.3.

QUINTA

Se implementó el modelo utilizando la técnica de regresión lineal aplicando un filtro de instancias no supervisadas, logrando obtener un modelo predictivo para cada indicador hospitalario del centro quirúrgico, es decir 12 ecuaciones; los cuales fueron obtenidos automáticamente mediante el software Weka. Ver punto 4.4.

SEXTA

Se logró evaluar el modelo propuesto mediante pruebas de precisión, con los datos del hospital y los datos del modelo para lo cual se aplicaron las ecuaciones obtenidas con la técnica de regresión lineal, obteniéndose un

83.9% de precisión, con lo cual se concluye que el modelo es válido. Ver punto 4.5.



SUGERENCIAS

PRIMERA

Adaptar el modelo predictivo de minería de datos para proyectos futuros de gestión hospitalaria, considerando el desarrollo de un software para la recolección de información y la automatización de reportes para el apoyo en la toma de decisiones.

SEGUNDA

Se sugiere utilizar técnicas de minería de datos para la mejora en la toma de decisión de los diferentes centros de atención en otras instituciones hospitalarias para hacer un comparativo y adoptar mejores procesos.

TERCERA

Para los nuevos investigadores se sugiere integrar al software a otras aplicaciones utilizadas en el hospital y extenderlo para su uso con otros dispositivos con mayor accesibilidad como: escáner, pantallas táctiles y otros.

CUARTA

Completar a la totalidad de indicadores hospitalarios para centros quirúrgicos (12) y realizar el procesamiento con mayor cantidad de informes operatorios, de esta manera también se podrán realizar más pruebas de modelado.

QUINTA

A los nuevos investigadores se sugiere aplicar otras técnicas de minería de datos y establecer un método de la implementación mediante el software Weka, también realizar pruebas filtrando indicadores y con la posibilidad de crear nuevas variables.

SEXTA

Realizar comparaciones de técnicas de minería de datos y evaluarlas con otras medidas de evaluación, de esta manera se lograrán obtener mejores resultados.



BIBLIOGRAFIA

- Barrios, M. (2010). *Modelo del Negocio*. Americana.
- Campoy Aranda , T., & Gomes Araújo, E. (2009). *Técnicas e instrumentos Cualitativos de Recogida de Datos*. España: EOS.
- Folgueiras Bertomeu, P. (2010). *Métodos y técnicas de recogida y análisis de información cualitativa*. Argentina: Universidad de Barcelona.
- Rodríguez Rojas, O. (2012). *Metodología para el Desarrollo de Proyectos en Minería de Datos CRISP-DM*. San Pedro, Costa Rica.
- Alejo Huarachi, A. M. (2009). *Portal Web de noticias del Perú aplicando tecnicas de recuperacion de información y mineria de datos*. Puno: Universidad Nacional del Altiplano.
- Aler, R. (2009). Tutorial Weka 3.6.0.
- Amorocho, G. (2015). *Mineria de datos: que es y 5 consejos para aprovecharla en las grandes empresas*. Obtenido de www.reportedigital.com
- Booch, Grady; Rumbaugh, jim; Jacobson, Ivar. (2013). *LenguajeUnificado de Modelado*. Obtenido de <http://elvex.ugr.es/decsai/java/pdf/3E-UML.pdf>
- Britos , P. (2010). *Objetivos de Negocio y Procesos de Minería de Datos Basados en Sistemas Inteligentes*. Argentina.
- Calderón Méndez, N. d. (2006). *Minería de datos una Herramienta para la Toma de Decisiones*. Guatemala.
- Calleja Gómez, A. (2010). *Mineria de datos para la Predicción del Precio de Automóviles de Segunda Mano*. Valencia, España.
- Carlos Marquéz Vera, C. R. (2012). *Predicción del Fracaso Escolar mediante Técnicas de Minería de Datos*.
- Castillo Rojas, W. A. (2013). *Inteligencia de Negocios. Minería de Datos III*. Chile.
- CEI. (2010). "Manual Básico para Elaborar Plan de Negocio para PYMEs". Nicaragua.
- Chapman, P. (20 de 09 de 2014). *Step-by-step Data Mining Guide*. Obtenido de www.crisp-dm.org/CRISPWP-0800.pdf

- Córdoba Fallas, L. (06 de Junio de 2011). *Minería de Datos* . Obtenido de Weka: <http://cor-mineriadedatos.blogspot.com/2011/06/weka.html>
- Cruz, R. S. (2015). *Técnicas y Herramientas de Minería de Datos*. España.
- Devore, J. (2008). *Probabilidad y Estadística para Ingenierías y Ciencias*. México: Cengage Learning Editores.
- Eduardo, A. (2015). *Definición General del Metodo*. Obtenido de <https://es.scribd.com/doc/53258812/DEFINICION-GENERAL-DEL-METODO>
- Ferré Grau, Xavier; Sánchez Segura, María Isabel. (2011). *Desarrollo Orientado a Objetos UML*. México.
- Freedman, D. (2009). *Statistical Models: Theory and Practice*. Phantom Fame.
- García González, F. J. (2013). *Aplicación de técnicas de Minería de Datos a datos obtenidos por el Centro Andaluz de Medio Ambiente (CEAMA)*. Granada: Universidad de Granada.
- García Jiménez, M., & Álvarez Sierra, A. (2013). *Análisis de Datos en WEKA – Pruebas de Selectividad*. Obtenido de <http://www.it.uc3m.es/jvillena/irc/practicas/06-07/28.pdf>
- Garrido, P. M. (Noviembre de 2012). *Búsqueda de patrones en Bases de Datos y su aplicación en las PYMES*. Veracruz, Mexico.
- Godoy, D. (2015). *Minería de datos Web*. Argentina.
- Gomez, A. J. (Diciembre de 2010). *Minería de Datos con Weka para la Predicción del Precio de Automóviles*. Valencia, Universidad Politécnica de Valencia, España.
- Gómez Díaz, Haydeé; Cerón Reyes, María de los Angeles. (2010). *Minería de datos*.
- Gonzales, F. J. (2013). *Aplicación de técnicas de minería de datos a datos obtenidos por el Centro Andaluz de Medio Ambiente*. Granada: Universidad de Granada.
- Gonzales, P. (2010). *Error Cuadrático Medio de Predicción para Modelos Estructurales de Series Temporales*. España: Estadística Española.
- Hasperué, W. (Febrero de 2012). *Extracción de Conocimiento en Grandes Bases de Datos Utilizando Estrategias Adaptativas*. La plata.
- Hasperué, W. (Febrero de 2012). *Extracción de Conocimiento en Grandes Bases de Datos Utilizando Estrategias Adaptativas*. La Plata, Argentina.

- Hernandez Orallo, José; Ramirez Quintana, María José; Ferri Ramirez, Cesar;. (2004). *Introduccion a la Minería de Datos*. Editorial Alhambra S. A. (SP).
- Hernández Gómez, G. (2014). Variables Cualitativas, Cuantitativas(Ordinales, Continuas y Discretas). Zaragoza, España.
- Hernández Orallo, E. (2011). *Diagramas UML*. Obtenido de http://www.teatroabadia.com/es/uploads/documentos/iagramas_del_uml.pdf
- Hernández Orallo, J., & Ferri Ramírez, C. (2013). Extracción Automática de Conocimiento en Bases de Datos e Ingeniería del Software.
- Hospital Manuel Núñez Butrón. (2013). Informes Operatorios. Puno, Perú: Archivo Central.
- Hospital Regional Manuel Núñez Butrón. (27 de 05 de 2014). *Hospital Regional Manuel Núñez Butrón*. Recuperado el 27 de 05 de 2015, de Hospital Regional Manuel Núñez Butrón: <http://hospitalregional-mnbpuno.gob.pe/images/ROF%20HRMNB.pdf>
- IBM. (2012). Manual CRISP-DM de IBM SPSS Modeler. Estados Unidos.
- IBM. (2012). *Manual CRISP-DM de IBM SPSS*.
- Jankovic, F. (2011). *Procesos científicos: predecir, interpretar datos, controlar variables*. Estados Unidos.
- Jímenez Galindo, Álvaro; Álvarez García, Hugo. (2010). Minería de Datos en la Educación. Madrid, España.
- Liz, Y. (2012). *Árboles de Decisiones*. Sevilla.
- Llavona Arregui, J. L. (2010). Terminología de Estadística y Minería de Datos en Lengua Inglesa. Madrid, España: ISBN.
- López Gaona, A. (2011). *El Modelo Entidad - Relación*. México: UNAM.
- Maroto , A., Boqué , R., Riu, J., & Rius, X. (2011). *INCERTIDUMBRE Y PRECISIÓN*. España.
- Mendoza Castillo, Sandra Jaqueline; Zavaleta Henriquez, Fernando. (Diciembre de 2013). Desarrollo de un Modelo de Minería de Datos para la Toma de Decisiones en la Gestión de Inventarios en Empresas Comerciales. Trujillo, Perú.

Mendoza Lara, L., Archiraico Agudo, W., Cuadros Oria, J., Oviedo Rodriguez, C., Avila Guerra, C., & Aparcana Carrasco, M. (Agosto de 2013). *Indicadores de Gestión y Evaluación Hospitalaria, Para Hospitales, Institutos y Diresa*. Lima, Perú.

Meredith, W. (2015). *Modelo de Implementación y Soporte a Sistemas de Oficinas de P.I.*

Microsoft. (2014). *Modelos de minería de datos (Analysis Services - Minería de datos)*. Obtenido de [https://msdn.microsoft.com/es-es/library/cc645779\(v=sql.120\).aspx#bkmk_mdldefine](https://msdn.microsoft.com/es-es/library/cc645779(v=sql.120).aspx#bkmk_mdldefine)

MINSA. (2013). *Indicadores Hospitalarios*. 67.

Moine, J. M., Haedo, A. S., & Gordillo, S. (2012). *Estudio Corporativo de los Datos*. Argentina: UTN Rosario.

Moine, J. M. (Abril de 2013). *Metodologías para el descubrimiento de Conocimiento en Bases de Datos*. La Plata, Argentina.

Moine, J. M. (2013). *Metodologías para el descubrimiento de conocimiento en bases de datos: un estudio comparativo*. Plata: Universidad Nacional de la Plata.

Moine, J. M. (2013). *Metodologías para el descubrimiento de conocimiento en bases de datos: un estudio comparativo*. La Plata: Universidad Nacional de La Plata.

Moine, Juan Miguel; Haedo, Ana; Gordillo, Silvia. (2011). *Estudio comparativo de metodologías para minería de datos*. Buenos Aires, Argentina.

Molero Castillo, G. G. (2008). *Desarrollo de un Modelo basado en técnicas de Minería de Datos para clasificar zona climatológicamente similares en el estado de Michoacán*. México: Universidad Nacional Autónoma de México.

Molina Felix, L. C. (Octubre de 2014). *Datamining: Torturando los Datos Hasta que Confiesen. Construyamos mejor software sin fronteras* (págs. http://sg.com.mx/sgvirtual/7/sessions/data-mining-torturando-los-datos-hasta-que-confiesen#.VYsxePl_Okp). México: SG Virtual Conference.

Molina López, J. M. (2006). *Técnicas de Análisis de Datos. Aplicaciones Practicas Utilizando Microsoft Excel y Weka*. España.

Mondragón, J. (03 de 12 de 2012). *Minería de Textos y el análisis predictivo. Uso de la minería de textos para incrementar el desempeño de los*. México.

- Monje Álvarez, C. A. (2011). *Metodología de la Investigación Cuantitativa y Cualitativa Guia didáctica*. Neiva.
- Moral Peláez, I. (2012). *Modelos de regresión: Lineal Simple y Regresión Logística*. España.
- Morone, G. (2013). *Métodos y Técnicas de la Investigación*. Obtenido de http://biblioteca.ucv.cl/site/servicios/documentos/metodologias_investigacion.pdf
- Párraga Álava, Jorge Antonio ; Alcívar Cevallos, Robert Abel. (2012). *Implementación de algoritmo a priori para la obtención de reglas de asociación en dataset sobre hepatitis*. Chile.
- Peralta Cochancela, D. (2009). *Proyecto de Minería de Datos para el Análisis del Comportamiento de los Clientes de Telecomunicaciones*. Cuenca: Universidad Politecnica Salesiana Sede Cuenca.
- Peralta, F. (2014). *Proceso de Conceptualización del Entendimiento del Negocio para Proyectos de Explotación de Información*. Buenos Aires: UTN.BA. Escuela de Posgrado.
- Ramos Román, I., & Dolado Cosín, J. J. (2007). *Técnicas cuantitativas para la gestión en la ingeniería del software*.
- Reyes, R. G. (2012). *Minería de datos para la toma de decisiones e inteligencia de negocios: Aplicaciones en la mercadotecnia*. Mexico.
- Rigeiro. (2012). *Procesamiento y análisis de los Datos*.
- Rodríguez Rojas, O. (Marzo de 2015). *Metodos Predictivos en Minería de Datos - PERSON*. Costa Rica.
- Rodríguez, H. (24 de Junio de 2009). *Gestión de Contenidos en Ciencia de la Información*. Obtenido de <https://ecibjaveriana.wordpress.com/2009/06/24/sobre-la-mineria-de-datos/>
- Roegiers, X., & Peyser, A. (2010). *El concepto y utilidad de la situación-problema*. BIEF.
- Rojo Abuín, J. M. (2017). *Regresión lineal múltiple*. Madrid.
- Romeau Guallart, P. (2010). *Minería de Datos Aplicada Al Análisis del Tratamiento Informativo de la Drogadicción*. Moncada: CEU.

Sandra Jaqueline Mendoza Castillo, Fernando Zavaleta Henrriquez. (2013). Desarrollo de un Modelo de Minería de Datos para la Toma de Decisiones en la Gestión de Inventarios en Empresas Comerciales. Trujillo, Perú.

Systems, P. S. (2011). *Modelado de Sistemas con UML*. Obtenido de <http://es.tldp.org/Tutoriales/doc-modelado-sistemas-UML/doc-modelado-sistemas-uml.pdf>

Torre de la Vega. (2012). *TEORÍA DE ERRORES*. Obtenido de <http://ocw.unican.es/enseanzas-tecnicas/fisica-i/practicas-1/Teoria%20de%20errores.pdf>

Turmero, I. (Marzo de 2011). Minería de Datos. *El Arte de Sacar Conocimiento de Grandes Volúmenes de Datos*. Puerto Ordaz, Guyana.

UCLM. (2012). *Guía práctica para la realización de medidas y el cálculo de errores*. Obtenido de https://www.uclm.es/profesorado/jmcolino/Docencia_archivos/Apuntes%20de%20C3%A1lculo%20de%20Errores.pdf

uclm. (2012). *Guía práctica para la realización de medidas y el cálculo de errores* . Obtenido de https://www.uclm.es/profesorado/jmcolino/Docencia_archivos/Apuntes%20de%20C3%A1lculo%20de%20Errores.pdf

UNED. (Febrero de 2012). Minería de Datos.

Universidad Autónoma del Estado de México . (2010). *Error absoluto y error relativo* . Mexico.

Vararey. (2015). *COEFICIENTE DE CORRELACIÓN LINEAL DE PEARSON* . Obtenido de <http://personal.us.es/vararey/adatos2/correlacion.pdf>

Vizcaino Garzon, P. A. (2008). *Aplicación de Técnicas de Inducción de Árboles de Decisión a Problemas de Clasificación mediante el uso de WEKA*. Bogotá.

Zapata, S. (Marzo de 2011). *Técnicas de Minería de Datos Basadas en Aprendizaje Automático*. Obtenido de <https://santiagozapatakdd.files.wordpress.com/2011/03/cursos-kdd-full-cap-3.pdf>

Zapata, S. (2011). *Técnicas de Minería de Datos basadas en Aprendizaje Automático* .

Zegarra Ciprés, Mercedes; Bou Llusar, Juan Carlos. (2004). Concepto y tipos de Dimensiones del Conocimiento: Configuración del Conocimiento Estrategico. España.





ANEXO 1: ENCUESTA

ENCUESTA

Fecha:01/02/2014

Responsable de la información:

Unidad de Estadística () Departamento de Cirugía () Centro Quirúrgico ()

Cirugía A () Cirugía B () Otro (Especificar).....

Se tienen datos históricos de pacientes atendidos en el área de cirugía?

Si()

En donde se almacenan
.....
.....

No()

Cuál es la cantidad de Personal con que cuenta el área de Cirugía?
.....
.....
.....

Que plantea Usted que podría ayudar en la mejora en la Toma de decisiones del Departamento de Cirugía.

.....
.....
.....
.....
.....
.....
.....

Que información sería importante predecir para la mejora en la toma de decisiones del Departamento de Cirugía

.....
.....
.....
.....
.....

Han elaborado algún indicador para la mejora en la toma de decisiones del Departamento de Cirugía.

Si()

Especifique

.....
.....
.....
.....
.....
.....
.....
.....
.....
.....

No()

Porque.....

Cuáles son los objetivos principales del departamento de Cirugía?

.....
.....
.....
.....
.....
.....

Gracias



ANEXO 2: FICHA DE OBSERVACIÓN

| Indicadores Hospitalarios | Indicadores | | | | Observaciones |
|---|-------------|---------------------------|----------------------|-------------------------|---------------|
| | Precisión | Coficiente de Correlación | Error medio absoluto | Error absoluto relativo | |
| Número de intervenciones quirúrgicas. | | | | | |
| Número de pacientes operados. | | | | | |
| Aprovechamiento de salas de operaciones. | | | | | |
| Grado de cumplimiento de las operaciones programadas. | | | | | |
| Rendimiento quirófono. | | | | | |
| Porcentaje de intervenciones quirúrgicas suspendidas. | | | | | |
| Promedio de días preoperatorios. | | | | | |
| Porcentaje de complicaciones intraoperatorias. | | | | | |
| Porcentaje de complicaciones postoperatorias. | | | | | |
| Porcentaje de complicaciones anestésicas. | | | | | |
| Porcentaje de complicaciones en recuperación. | | | | | |
| Tasa de Mortalidad | | | | | |
| | | | | | |

ANEXO 3: Obtención de Resultados Actuales de Indicadores Hospitalarios

| INDICADORES | % DE PRECISIÓN ACTUAL |
|--|-----------------------|
| Número de Intervenciones Quirúrgicas | |
| Número de Pacientes Operados | |
| Aprovechamiento Sala de Operaciones | |
| Rendimiento de Quirófano | |
| Porcentaje de Intervenciones Quirúrgicas Suspendidas | |
| Tasa de mortalidad | |





ANEXO 4 : INFORME OPERATORIO

000778



HOSPITAL REGIONAL "MANUEL NÚÑEZ BUTRON" PUNO

INFORME OPERATORIO



INFORME N° _____ FECHA DE INSCRIPCIÓN 1 / 1 N° de H. CLINICA: _____

| | | | | | |
|--|------------|---|------------------------------|-------------------------------------|--|
| APELLIDOS | PATERNO | MATERNO | NOMBRES | EDAD | SEXO |
| | Tapia | Chuspa | Estela | 18 | F |
| OCUPACIÓN | Estudiante | | PROCEDENCIA | Puno | |
| DEPARTAMENTO | Cuzco | SERVICIO | CA | PISO | SALA N° CAMA N° 9 |
| Días de Internamiento Antes de Operación: 4 días | | | Interv. programada | <input checked="" type="checkbox"/> | Interv. de Emergencia <input type="checkbox"/> |
| DIAGNOSTICO PRE OPERATORIO: Absceso de Paredi HI, Desecho, osteítis de | | | PRE MEDICACIÓN: Tetraciclina | | |
| ANESTESIA Local | | | ANESTESICO | CANTIDAD | |
| ANESTESIOLOGO NOMBRE: Dr. Dora | | | FIRMA | | |
| INTERVENCIÓN 1: Punción Quirúrgica y Sutura | | | CÓDIGO | | |
| INTERVENCIÓN 2: | | | CÓDIGO | | |
| INTERVENCIÓN 3: | | | CÓDIGO | | |
| Hora de Ingreso de Paciente: 9 AM | | Sala de Operación N°: 05 | | Hora de Egreso: | |
| Fecha de Operación: 15 de Julio | | del 2013; Hora de inicio: 9:30 | | Hora de Terminó: 9:45 Duración | |
| CIRUJANO NOMBRE: Dr. Gomez | | | FIRMA: | | |
| PRIMER AYUDANTE: Dr. Rg Hector Vilca Cordero | | | FIRMA: | | |
| SEGUNDO AYUDANTE: | | | FIRMA: | | |
| INSTRUMENTISTA NOMBRE: | | | FIRMA: | | |
| CIRCULANTE NOMBRE: Dra Maria Rojas P. | | | FIRMA: | | |
| PROTOCOLO OPERATORIO: Se realiza inspección y xibografía 2% de a mi del do amitar en puerca, edroegria, quia AA, CC, se realiza cura y luego se procede a suturar heridas. | | | | | |
| DIAGNOSTICO POST OPERATORIO: Curación | | | | | |
| SALA DE RECUPERACIÓN | | Hora de Ingreso | | Hora de Egreso | |
| Nombre: | | | | | |
| COMPLICACIONES POR ANESTESIA | | NO <input type="checkbox"/> SI <input type="checkbox"/> Especifique: | | | |
| EN ACTO QUIRURGICO | | NO <input type="checkbox"/> SI <input type="checkbox"/> Especifique: | | | |
| EN SALA DE RECUPERACIÓN | | NO <input type="checkbox"/> SI <input type="checkbox"/> Especifique: | | | |
| CONDICIÓN DE EGRESO | | <input type="checkbox"/> Por Anestesia <input type="checkbox"/> En Operación <input type="checkbox"/> En Recuperación SUSPENDIDO <input type="checkbox"/> Especifique: | | | |
| VIVO <input type="checkbox"/> MUERTO <input type="checkbox"/> | | MUESTRA ENVIADA A PATOLOGIA: NO <input type="checkbox"/> SI <input type="checkbox"/> Especifique: | | | |
| OBSERVACIONES: | | | | | |
| FIRMA | | | | | |

ANEXO 5: CONSULTAS SQL

- **Número de intervenciones quirúrgicas (NIQ):**

| Anio | Mes | Intervenciones |
|------|-----|----------------|
| 2012 | 5 | 105 |
| 2012 | 6 | 41 |
| 2012 | 7 | 151 |
| 2012 | 8 | 132 |
| 2012 | 9 | 98 |
| 2012 | 10 | 88 |
| 2012 | 11 | 36 |
| 2012 | 12 | 118 |
| 2013 | 1 | 117 |
| 2013 | 2 | 147 |
| 2013 | 3 | 153 |
| 2013 | 4 | 159 |
| 2013 | 5 | 20 |
| 2013 | 12 | 1 |

- **Número de pacientes operados (NPO):**

| Anio | Mes | Pacientes |
|------|-----|-----------|
| 2012 | 5 | 105 |
| 2012 | 6 | 40 |
| 2012 | 7 | 149 |
| 2012 | 8 | 132 |
| 2012 | 9 | 96 |
| 2012 | 10 | 85 |
| 2012 | 11 | 36 |
| 2012 | 12 | 119 |
| 2013 | 1 | 114 |
| 2013 | 2 | 143 |
| 2013 | 3 | 153 |
| 2013 | 4 | 155 |
| 2013 | 5 | 51 |
| 2013 | 12 | 1 |

- **Aprovechamiento de salas de operaciones (ASO):**

| Anio | Mes | suma |
|------|-----|--------|
| 2012 | 5 | 0.3042 |
| 2012 | 6 | 0.1292 |
| 2012 | 7 | 0.3778 |
| 2012 | 8 | 0.3958 |
| 2012 | 9 | 0.2917 |
| 2012 | 10 | 0.2944 |
| 2012 | 11 | 0.0847 |
| 2012 | 12 | 0.2833 |
| 2013 | 1 | 0.3417 |
| 2013 | 2 | 0.325 |
| 2013 | 3 | 0.4528 |
| 2013 | 4 | 0.4667 |
| 2013 | 5 | 0.0917 |
| 2013 | 12 | 0.0125 |

- **Grado de cumplimiento de las operaciones programadas (GCOP):**

| Anio | Mes | Programadas | Suspendidos | Grado |
|------|-----|-------------|-------------|-------|
| 2012 | 5 | 108 | 3 | 97.22 |
| 2012 | 6 | 41 | 0 | 100 |
| 2012 | 7 | 151 | 0 | 100 |
| 2012 | 8 | 132 | 0 | 100 |
| 2012 | 9 | 98 | 0 | 100 |
| 2012 | 10 | 88 | 0 | 100 |
| 2012 | 11 | 36 | 0 | 100 |
| 2012 | 12 | 120 | 2 | 98.33 |
| 2013 | 1 | 117 | 0 | 100 |
| 2013 | 2 | 147 | 0 | 100 |
| 2013 | 3 | 154 | 1 | 99.35 |
| 2013 | 4 | 159 | 0 | 100 |
| 2013 | 5 | 52 | 32 | 38.46 |
| 2013 | 12 | 1 | 0 | 100 |

- Rendimiento quirófono (RQ):

| Anio | Mes | Promedio |
|------|-----|----------|
| 2012 | 5 | 21 |
| 2012 | 6 | 8.2 |
| 2012 | 7 | 30.2 |
| 2012 | 8 | 26.4 |
| 2012 | 9 | 19.6 |
| 2012 | 10 | 17.6 |
| 2012 | 11 | 7.2 |
| 2012 | 12 | 23.6 |
| 2013 | 1 | 23.4 |
| 2013 | 2 | 29.4 |
| 2013 | 3 | 30.6 |
| 2013 | 4 | 31.8 |
| 2013 | 5 | 4 |
| 2013 | 12 | 0.2 |

- Porcentaje de intervenciones quirúrgicas suspendidas (PIQS)

| Anio | Mes | Porcentaje |
|------|-----|------------|
| 2012 | 5 | 2.78 |
| 2012 | 6 | 0 |
| 2012 | 7 | 0 |
| 2012 | 8 | 0 |
| 2012 | 9 | 0 |
| 2012 | 10 | 0 |
| 2012 | 11 | 0 |
| 2012 | 12 | 1.67 |
| 2013 | 1 | 0 |
| 2013 | 2 | 0 |
| 2013 | 3 | 0.65 |
| 2013 | 4 | 0 |
| 2013 | 5 | 61.54 |
| 2013 | 12 | 0 |
| 2014 | 12 | 0 |

- Promedio de días preoperatorios.

| Anio | Mes | Promedio |
|------|-----|----------|
| 2012 | 5 | 1.38 |
| 2012 | 6 | 0.68 |
| 2012 | 7 | 0.85 |
| 2012 | 8 | 1.08 |
| 2012 | 9 | 1.43 |
| 2012 | 10 | 1.15 |
| 2012 | 11 | 0.89 |
| 2012 | 12 | 1 |
| 2013 | 1 | 1.03 |
| 2013 | 2 | 1.13 |
| 2013 | 3 | 1.22 |
| 2013 | 4 | 1.33 |
| 2013 | 5 | 1.3 |
| 2013 | 12 | 1 |

- Porcentaje de infecciones postquirúrgicas (PIPQ):

| Anio | Mes | PIPQ |
|------|-----|--------|
| 2012 | 5 | 27.619 |
| 2012 | 6 | 12.195 |
| 2012 | 7 | 52.318 |
| 2012 | 8 | 100 |
| 2012 | 9 | 100 |
| 2012 | 10 | 88.636 |
| 2012 | 11 | 0 |
| 2012 | 12 | 22.034 |
| 2013 | 1 | 5.128 |
| 2013 | 2 | 12.245 |
| 2013 | 3 | 2.614 |
| 2013 | 4 | 1.258 |
| 2013 | 5 | 0 |
| 2013 | 12 | 0 |

• **Porcentaje de complicaciones intraoperatorias (PCI):**

| Anio | Mes | Promedio |
|------|-----|----------|
| 2012 | 5 | 0 |
| 2012 | 6 | 0 |
| 2012 | 7 | 0 |
| 2012 | 8 | 0 |
| 2012 | 9 | 0 |
| 2012 | 10 | 0 |
| 2012 | 11 | 0 |
| 2012 | 12 | 0 |
| 2013 | 1 | 0 |
| 2013 | 2 | 0 |
| 2013 | 3 | 0 |
| 2013 | 4 | 0.01 |
| 2013 | 5 | 0 |
| 2013 | 12 | 1 |

• **Porcentaje de complicaciones anestésicas (PCA):**

| Anio | Mes | Promedio |
|------|-----|----------|
| 2012 | 5 | 0 |
| 2012 | 6 | 0 |
| 2012 | 7 | 0 |
| 2012 | 8 | 0 |
| 2012 | 9 | 0 |
| 2012 | 10 | 0.02 |
| 2012 | 11 | 0 |
| 2012 | 12 | 0 |
| 2013 | 1 | 0 |
| 2013 | 2 | 0 |
| 2013 | 3 | 0 |
| 2013 | 4 | 0 |
| 2013 | 5 | 0 |
| 2013 | 12 | 1 |

- **Porcentaje de complicaciones en recuperación.**

| Anio | Mes | Promedio |
|------|-----|----------|
| 2012 | 5 | 0 |
| 2012 | 6 | 0 |
| 2012 | 7 | 0 |
| 2012 | 8 | 0 |
| 2012 | 9 | 0 |
| 2012 | 10 | 0.02 |
| 2012 | 11 | 0 |
| 2012 | 12 | 0 |
| 2013 | 1 | 0 |
| 2013 | 2 | 0 |
| 2013 | 3 | 0 |
| 2013 | 4 | 0 |
| 2013 | 5 | 0 |
| 2013 | 12 | 0 |

- **Tasa de mortalidad en el centro quirúrgico (TM):**

| Anio | Mes | Promedio |
|------|-----|----------|
| 2012 | 5 | 0 |
| 2012 | 6 | 0 |
| 2012 | 7 | 0 |
| 2012 | 8 | 0 |
| 2012 | 9 | 0 |
| 2012 | 10 | 1.14 |
| 2012 | 11 | 0 |
| 2012 | 12 | 0 |
| 2013 | 1 | 0.85 |
| 2013 | 2 | 0 |
| 2013 | 3 | 0 |
| 2013 | 4 | 1.26 |
| 2013 | 5 | 0 |
| 2013 | 12 | 0 |

• Resumen indicadores por día:

| Año | Mes | Día | NIQ | NPO | ASO | GCOP | RQ | PIQS | PDP | PIPQ | PCI | PCA | PCR | TM |
|------|-----|-----|-----|-----|--------|--------|-----|--------|-------|--------|-----|-----|-----|----|
| 2012 | 5 | 11 | 1 | 1 | 0.833 | 100 | 0.2 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 5 | 12 | 4 | 4 | 10 | 100 | 0.8 | 0 | 12.25 | 50 | 0 | 0 | 0 | 0 |
| 2012 | 5 | 13 | 3 | 3 | 2.5 | 100 | 0.6 | 0 | 0 | 33.333 | 0 | 0 | 0 | 0 |
| 2012 | 5 | 14 | 6 | 7 | 12.5 | 85.714 | 1.2 | 14.286 | 0.714 | 66.667 | 0 | 0 | 0 | 0 |
| 2012 | 5 | 15 | 3 | 3 | 5 | 100 | 0.6 | 0 | 1.333 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 5 | 16 | 4 | 4 | 3.333 | 100 | 0.8 | 0 | 0.25 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 5 | 17 | 7 | 7 | 9.167 | 100 | 1.4 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 5 | 18 | 6 | 6 | 10.833 | 100 | 1.2 | 0 | 0.333 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 5 | 19 | 3 | 3 | 3.333 | 100 | 0.6 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 5 | 20 | 2 | 2 | 6.667 | 100 | 0.4 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 5 | 21 | 5 | 5 | 10 | 100 | 1 | 0 | 2.6 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 5 | 22 | 6 | 7 | 11.667 | 85.714 | 1.2 | 14.286 | 0.286 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 5 | 23 | 4 | 4 | 4.167 | 100 | 0.8 | 0 | 0.5 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 5 | 24 | 5 | 6 | 5.833 | 83.333 | 1 | 16.667 | 2.333 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 5 | 25 | 7 | 7 | 9.167 | 100 | 1.4 | 0 | 0.714 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 5 | 26 | 6 | 6 | 5.833 | 100 | 1.2 | 0 | 0.333 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 5 | 27 | 3 | 3 | 2.5 | 100 | 0.6 | 0 | 0 | 66.667 | 0 | 0 | 0 | 0 |
| 2012 | 5 | 28 | 5 | 5 | 5.833 | 100 | 1 | 0 | 0.8 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 5 | 29 | 9 | 9 | 11.667 | 100 | 1.8 | 0 | 1.333 | 33.333 | 0 | 0 | 0 | 0 |
| 2012 | 5 | 30 | 6 | 6 | 23.333 | 100 | 1.2 | 0 | 2.167 | 66.667 | 0 | 0 | 0 | 0 |
| 2012 | 5 | 31 | 10 | 10 | 22.5 | 100 | 2 | 0 | 0.2 | 70 | 0 | 0 | 0 | 0 |
| 2012 | 6 | 1 | 5 | 5 | 7.5 | 100 | 1 | 0 | 0 | 20 | 0 | 0 | 0 | 0 |
| 2012 | 6 | 2 | 3 | 3 | 19.167 | 100 | 0.6 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 6 | 4 | 3 | 3 | 2.5 | 100 | 0.6 | 0 | 0.667 | 0 | 0 | 0 | 0 | 0 |

| Año | Mes | Día | NIQ | NPO | ASO | GCOP | RQ | PIQS | PDP | PIPQ | PCI | PCA | PCR | TM |
|------|-----|-----|-----|-----|--------|------|-----|------|-------|--------|-----|-----|-----|----|
| 2012 | 6 | 5 | 1 | 1 | 0.833 | 100 | 0.2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 6 | 6 | 5 | 5 | 5 | 100 | 1 | 0 | 0.4 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 6 | 7 | 6 | 6 | 8.333 | 100 | 1.2 | 0 | 0.833 | 16.667 | 0 | 0 | 0 | 0 |
| 2012 | 6 | 8 | 4 | 4 | 15 | 100 | 0.8 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 6 | 9 | 1 | 1 | 0.833 | 100 | 0.2 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 6 | 10 | 2 | 2 | 2.5 | 100 | 0.4 | 0 | 3.5 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 6 | 11 | 5 | 5 | 5 | 100 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 6 | 12 | 3 | 3 | 4.167 | 100 | 0.6 | 0 | 0.667 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 6 | 29 | 1 | 1 | 0.833 | 100 | 0.2 | 0 | 2 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 6 | 30 | 2 | 2 | 1.667 | 100 | 0.4 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 2 | 7 | 7 | 10 | 100 | 1.4 | 0 | 1.286 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 3 | 3 | 3 | 4.167 | 100 | 0.6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 4 | 2 | 2 | 0.833 | 100 | 0.4 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 5 | 4 | 4 | 4.167 | 100 | 0.8 | 0 | 2.75 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 6 | 6 | 6 | 5.833 | 100 | 1.2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 7 | 6 | 6 | 13.333 | 100 | 1.2 | 0 | 0.333 | 16.667 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 8 | 2 | 2 | 4.167 | 100 | 0.4 | 0 | 1.5 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 9 | 6 | 6 | 5.833 | 100 | 1.2 | 0 | 0.667 | 66.667 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 10 | 11 | 11 | 21.667 | 100 | 2.2 | 0 | 0.455 | 36.364 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 11 | 6 | 6 | 4.167 | 100 | 1.2 | 0 | 0.333 | 16.667 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 12 | 3 | 3 | 1.667 | 100 | 0.6 | 0 | 0.333 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 13 | 9 | 9 | 11.667 | 100 | 1.8 | 0 | 3 | 44.444 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 14 | 4 | 4 | 5 | 100 | 0.8 | 0 | 1.25 | 25 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 15 | 4 | 4 | 2.5 | 100 | 0.8 | 0 | 0.25 | 25 | 0 | 0 | 0 | 0 |

| Año | Mes | Día | NIQ | NPO | ASO | GCOP | RQ | PIQS | PDP | PIPQ | PCI | PCA | PCR | TM |
|------|-----|-----|-----|-----|--------|------|-----|------|-------|--------|-----|-----|-----|----|
| 2012 | 7 | 16 | 6 | 6 | 15 | 100 | 1.2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 17 | 5 | 5 | 5 | 100 | 1 | 0 | 0.4 | 20 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 18 | 6 | 6 | 8.333 | 100 | 1.2 | 0 | 0.833 | 16.667 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 19 | 8 | 8 | 22.5 | 100 | 1.6 | 0 | 0.75 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 20 | 8 | 8 | 9.167 | 100 | 1.6 | 0 | 1 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 21 | 4 | 4 | 16.667 | 100 | 0.8 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 22 | 2 | 2 | 1.667 | 100 | 0.4 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 23 | 5 | 5 | 5.833 | 100 | 1 | 0 | 1.2 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 24 | 5 | 5 | 5.833 | 100 | 1 | 0 | 2.4 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 25 | 5 | 5 | 4.167 | 100 | 1 | 0 | 0.6 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 26 | 6 | 6 | 4.167 | 100 | 1.2 | 0 | 0.667 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 27 | 1 | 1 | 0.833 | 100 | 0.2 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 28 | 4 | 4 | 4.167 | 100 | 0.8 | 0 | 1 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 29 | 4 | 4 | 5.833 | 100 | 0.8 | 0 | 0.75 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 30 | 3 | 3 | 2.5 | 100 | 0.6 | 0 | 0.333 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 7 | 31 | 6 | 6 | 7.5 | 100 | 1.2 | 0 | 0.333 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 1 | 2 | 2 | 4.167 | 100 | 0.4 | 0 | 2.5 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 2 | 5 | 5 | 10 | 100 | 1 | 0 | 2.2 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 3 | 7 | 7 | 9.167 | 100 | 1.4 | 0 | 1.714 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 4 | 3 | 3 | 5 | 100 | 0.6 | 0 | 1.333 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 6 | 3 | 3 | 5.833 | 100 | 0.6 | 0 | 2.333 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 7 | 6 | 6 | 8.333 | 100 | 1.2 | 0 | 1.833 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 8 | 2 | 2 | 2.5 | 100 | 0.4 | 0 | 2 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 9 | 6 | 6 | 37.5 | 100 | 1.2 | 0 | 1.667 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 10 | 1 | 1 | 4.167 | 100 | 0.2 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 11 | 3 | 3 | 3.333 | 100 | 0.6 | 0 | 1 | 100 | 0 | 0 | 0 | 0 |

| Año | Mes | Día | NIQ | NPO | ASO | GCOP | RQ | PIQS | PDP | PIPQ | PCI | PCA | PCR | TM |
|------|-----|-----|-----|-----|--------|------|-----|------|-------|------|-----|-----|-----|----|
| 2012 | 8 | 12 | 5 | 5 | 5 | 100 | 1 | 0 | 0.4 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 13 | 3 | 3 | 3.333 | 100 | 0.6 | 0 | 1.667 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 14 | 9 | 9 | 11.667 | 100 | 1.8 | 0 | 1.333 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 15 | 3 | 3 | 0.833 | 100 | 0.6 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 16 | 9 | 9 | 10 | 100 | 1.8 | 0 | 0.333 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 17 | 3 | 3 | 5 | 100 | 0.6 | 0 | 2 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 18 | 3 | 3 | 2.5 | 100 | 0.6 | 0 | 1.333 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 19 | 4 | 4 | 3.333 | 100 | 0.8 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 20 | 7 | 7 | 25 | 100 | 1.4 | 0 | 0.143 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 21 | 5 | 5 | 5.833 | 100 | 1 | 0 | 1 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 22 | 4 | 4 | 15 | 100 | 0.8 | 0 | 0.75 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 23 | 8 | 8 | 11.667 | 100 | 1.6 | 0 | 2.375 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 24 | 5 | 5 | 6.667 | 100 | 1 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 25 | 3 | 3 | 2.5 | 100 | 0.6 | 0 | 0.333 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 26 | 4 | 4 | 12.5 | 100 | 0.8 | 0 | 0.25 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 27 | 5 | 5 | 15.833 | 100 | 1 | 0 | 1 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 28 | 4 | 4 | 5.833 | 100 | 0.8 | 0 | 1.25 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 29 | 6 | 6 | 6.667 | 100 | 1.2 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 8 | 31 | 4 | 4 | 15.833 | 100 | 0.8 | 0 | 0.75 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 1 | 3 | 3 | 6.667 | 100 | 0.6 | 0 | 4 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 2 | 1 | 1 | 0.833 | 100 | 0.2 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 3 | 4 | 4 | 5.833 | 100 | 0.8 | 0 | 1.25 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 4 | 3 | 3 | 3.333 | 100 | 0.6 | 0 | 0.333 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 5 | 5 | 5 | 5 | 100 | 1 | 0 | 0.6 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 6 | 7 | 7 | 9.167 | 100 | 1.4 | 0 | 0.286 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 7 | 4 | 4 | 7.5 | 100 | 0.8 | 0 | 10.5 | 100 | 0 | 0 | 0 | 0 |

| Año | Mes | Día | NIQ | NPO | ASO | GCOP | RQ | PIQS | PDP | PIPQ | PCI | PCA | PCR | TM |
|------|-----|-----|-----|-----|--------|------|-----|------|-------|------|-----|-----|-----|----|
| 2012 | 9 | 8 | 1 | 1 | 0.833 | 100 | 0.2 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 9 | 2 | 2 | 1.667 | 100 | 0.4 | 0 | 1 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 10 | 5 | 5 | 7.5 | 100 | 1 | 0 | 0.4 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 11 | 7 | 7 | 20.833 | 100 | 1.4 | 0 | 3.714 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 12 | 6 | 6 | 13.333 | 100 | 1.2 | 0 | 1 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 13 | 6 | 6 | 7.5 | 100 | 1.2 | 0 | 0.833 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 14 | 3 | 3 | 4.167 | 100 | 0.6 | 0 | 0.667 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 15 | 1 | 1 | 0 | 100 | 0.2 | 0 | 1 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 16 | 3 | 3 | 5 | 100 | 0.6 | 0 | 1 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 17 | 6 | 6 | 25.833 | 100 | 1.2 | 0 | 1.333 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 18 | 6 | 6 | 7.5 | 100 | 1.2 | 0 | 0.5 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 19 | 2 | 2 | 1.667 | 100 | 0.4 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 20 | 2 | 2 | 2.5 | 100 | 0.4 | 0 | 1 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 21 | 1 | 1 | 0 | 100 | 0.2 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 22 | 1 | 1 | 0.833 | 100 | 0.2 | 0 | 1 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 24 | 3 | 3 | 1.667 | 100 | 0.6 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 25 | 3 | 3 | 3.333 | 100 | 0.6 | 0 | 0.333 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 26 | 3 | 3 | 1.667 | 100 | 0.6 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 27 | 4 | 4 | 5 | 100 | 0.8 | 0 | 3.25 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 28 | 2 | 2 | 1.667 | 100 | 0.4 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 9 | 29 | 4 | 4 | 14.167 | 100 | 0.8 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 10 | 1 | 1 | 1 | 1.667 | 100 | 0.2 | 0 | 3 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 10 | 2 | 4 | 4 | 5 | 100 | 0.8 | 0 | 1.75 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 10 | 3 | 1 | 1 | 0.833 | 100 | 0.2 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 10 | 4 | 5 | 5 | 6.667 | 100 | 1 | 0 | 0 | 100 | 0 | 0 | 0 | 0 |

| Año | Mes | Día | NIQ | NPO | ASO | GCOP | RQ | PIQS | PDP | PIPQ | PCI | PCA | PCR | TM |
|------|-----|-----|-----|-----|--------|------|-----|------|-----|-------|--------|-----|--------|--------|
| 2012 | 10 | 5 | 3 | 3 | 5 | 100 | 0.6 | | 0 | 0.333 | 100 | 0 | 0 | 0 |
| 2012 | 10 | 6 | 4 | 4 | 5.833 | 100 | 0.8 | | 0 | 0.5 | 100 | 0 | 0 | 0 |
| 2012 | 10 | 8 | 7 | 7 | 9.167 | 100 | 1.4 | | 0 | 0.286 | 100 | 0 | 0 | 0 |
| 2012 | 10 | 9 | 6 | 6 | 15 | 100 | 1.2 | | 0 | 0.667 | 100 | 0 | 0 | 0 |
| 2012 | 10 | 10 | 1 | 1 | 0.833 | 100 | 0.2 | | 0 | | 100 | 0 | 0 | 0 |
| 2012 | 10 | 11 | 2 | 2 | 1.667 | 100 | 0.4 | | 0 | 2 | 100 | 0 | 0 | 0 |
| 2012 | 10 | 12 | 1 | 1 | 0.833 | 100 | 0.2 | | 0 | 0 | 100 | 0 | 0 | 0 |
| 2012 | 10 | 13 | 3 | 3 | 2.5 | 100 | 0.6 | | 0 | 0.667 | 100 | 0 | 33.333 | 33.333 |
| 2012 | 10 | 14 | 1 | 1 | 0.833 | 100 | 0.2 | | 0 | 2 | 100 | 0 | 0 | 0 |
| 2012 | 10 | 15 | 3 | 3 | 5.833 | 100 | 0.6 | | 0 | 0.667 | 100 | 0 | 0 | 0 |
| 2012 | 10 | 16 | 4 | 4 | 10.833 | 100 | 0.8 | | 0 | 0.25 | 100 | 0 | 0 | 0 |
| 2012 | 10 | 17 | 7 | 7 | 7.5 | 100 | 1.4 | | 0 | 2 | 100 | 0 | 0 | 0 |
| 2012 | 10 | 18 | 5 | 5 | 16.667 | 100 | 1 | | 0 | 4 | 100 | 0 | 0 | 0 |
| 2012 | 10 | 19 | 1 | 1 | 0 | 100 | 0.2 | | 0 | 0 | 100 | 0 | 0 | 0 |
| 2012 | 10 | 20 | 1 | 1 | 0 | 100 | 0.2 | | 0 | 0 | 100 | 0 | 0 | 0 |
| 2012 | 10 | 21 | 1 | 1 | 0.833 | 100 | 0.2 | | 0 | 1 | 100 | 0 | 0 | 0 |
| 2012 | 10 | 22 | 6 | 5 | 5.833 | 100 | 1.2 | | 0 | 2 | 33.333 | 0 | 0 | 0 |
| 2012 | 10 | 23 | 3 | 3 | 5 | 100 | 0.6 | | 0 | 0.333 | 33.333 | 0 | 0 | 0 |
| 2012 | 10 | 24 | 5 | 5 | 8.333 | 100 | 1 | | 0 | 0 | 80 | 0 | 20 | 20 |
| 2012 | 10 | 25 | 5 | 5 | 37.5 | 100 | 1 | | 0 | 4 | 60 | 0 | 0 | 0 |
| 2012 | 10 | 26 | 5 | 5 | 16.667 | 100 | 1 | | 0 | 0.8 | 80 | 0 | 0 | 0 |
| 2012 | 10 | 27 | 3 | 3 | 3.333 | 100 | 0.6 | | 0 | 0.333 | 100 | 0 | 0 | 0 |
| 2012 | 11 | 22 | 9 | 9 | 13.333 | 100 | 1.8 | | 0 | 2.333 | 0 | 0 | 0 | 0 |
| 2012 | 11 | 23 | 4 | 4 | 5 | 100 | 0.8 | | 0 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 11 | 24 | 3 | 3 | 3.333 | 100 | 0.6 | | 0 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 11 | 25 | 4 | 4 | 5 | 100 | 0.8 | | 0 | 0.5 | 0 | 0 | 0 | 0 |



| Año | Mes | Día | NIQ | NPO | ASO | GCOP | RQ | PIQS | PDP | PIPQ | PCI | PCA | PCR | TM |
|------|-----|-----|-----|-----|--------|------|-----|------|-------|--------|-----|-----|-----|----|
| 2012 | 11 | 26 | 4 | 4 | 4.167 | 100 | 0.8 | 0 | 0.5 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 11 | 27 | 7 | 7 | 10.833 | 100 | 1.4 | 0 | 0.571 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 11 | 28 | 1 | 1 | 2.5 | 100 | 0.2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 11 | 29 | 2 | 2 | 2.5 | 100 | 0.4 | 0 | 0.5 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 11 | 30 | 2 | 2 | 2.5 | 100 | 0.4 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 1 | 1 | 1 | 0.833 | 100 | 0.2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 2 | 2 | 2 | 3.333 | 100 | 0.4 | 0 | 3 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 3 | 2 | 2 | 0.833 | 100 | 0.4 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 4 | 4 | 4 | 5.833 | 100 | 0.8 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 5 | 6 | 6 | 8.333 | 100 | 1.2 | 0 | 1.667 | 16.667 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 6 | 7 | 7 | 8.333 | 100 | 1.4 | 0 | 0.857 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 7 | 6 | 6 | 6.667 | 100 | 1.2 | 0 | 0.833 | 16.667 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 8 | 1 | 1 | 0 | 100 | 0.2 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 9 | 2 | 2 | 2.5 | 100 | 0.4 | 0 | 0.5 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 10 | 6 | 6 | 7.5 | 100 | 1.2 | 0 | 4.333 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 11 | 5 | 5 | 3.333 | 100 | 1 | 0 | 0 | 20 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 12 | 0 | 1 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 13 | 4 | 5 | 2.5 | 80 | 0.8 | 20 | 1.2 | 25 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 14 | 6 | 6 | 6.667 | 100 | 1.2 | 0 | 0.833 | 83.333 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 15 | 1 | 1 | 1.667 | 100 | 0.2 | 0 | 6 | 100 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 16 | 3 | 3 | 4.167 | 100 | 0.6 | 0 | 2.333 | 33.333 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 17 | 8 | 8 | 30.833 | 100 | 1.6 | 0 | 1.375 | 87.5 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 18 | 2 | 2 | 3.333 | 100 | 0.4 | 0 | 0.5 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 19 | 4 | 4 | 3.333 | 100 | 0.8 | 0 | 0.25 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 20 | 5 | 5 | 5 | 100 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

| Año | Mes | Día | NIQ | NPO | ASO | GCOP | RQ | PIQS | PDP | PIPQ | PCI | PCA | PCR | TM |
|------|-----|-----|-----|-----|--------|------|-----|------|-------|------|-----|-----|-----|------|
| 2012 | 12 | 21 | 7 | 7 | 8.333 | 100 | 1.4 | 0 | 0.571 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 22 | 6 | 6 | 7.5 | 100 | 1.2 | 0 | 0.167 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 24 | 7 | 7 | 7.5 | 100 | 1.4 | 0 | 0.571 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 25 | 4 | 4 | 4.167 | 100 | 0.8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 26 | 5 | 5 | 4.167 | 100 | 1 | 0 | 0.4 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 27 | 5 | 5 | 16.667 | 100 | 1 | 0 | 0.6 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 28 | 2 | 2 | 1.667 | 100 | 0.4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 29 | 3 | 3 | 2.5 | 100 | 0.6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 30 | 2 | 2 | 1.667 | 100 | 0.4 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2012 | 12 | 31 | 2 | 2 | 2.5 | 100 | 0.4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 1 | 2 | 2 | 1.667 | 100 | 0.4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 2 | 2 | 2 | 2.5 | 100 | 0.4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 3 | 5 | 5 | 18.333 | 100 | 1 | 0 | 0.2 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 4 | 2 | 2 | 1.667 | 100 | 0.4 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 5 | 2 | 2 | 4.167 | 100 | 0.4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 6 | 4 | 4 | 5 | 100 | 0.8 | 0 | 0.5 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 7 | 2 | 2 | 4.167 | 100 | 0.4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 8 | 7 | 7 | 12.5 | 100 | 1.4 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 9 | 4 | 4 | 5.833 | 100 | 0.8 | 0 | 0.5 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 10 | 4 | 4 | 1.667 | 100 | 0.8 | 0 | 0.25 | 50 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 11 | 8 | 8 | 15 | 100 | 1.6 | 0 | 0 | 12.5 | 0 | 0 | 0 | 12.5 |
| 2013 | 1 | 13 | 4 | 4 | 16.667 | 100 | 0.8 | 0 | 0.25 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 14 | 4 | 4 | 5 | 100 | 0.8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 15 | 6 | 6 | 16.667 | 100 | 1.2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 16 | 5 | 5 | 6.667 | 100 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |

| Año | Mes | Día | NIQ | NPO | ASO | GCOP | RQ | PIQS | PDP | PIPQ | PCI | PCA | PCR | TM |
|------|-----|-----|-----|-----|--------|------|-----|------|-------|--------|-----|-----|-----|----|
| 2013 | 1 | 17 | 4 | 4 | 4.167 | 100 | 0.8 | 0 | 7.75 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 18 | 1 | 1 | 0.833 | 100 | 0.2 | 0 | 4 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 20 | 1 | 1 | 0.833 | 100 | 0.2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 21 | 5 | 5 | 7.5 | 100 | 1 | 0 | 0.4 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 22 | 4 | 4 | 15.833 | 100 | 0.8 | 0 | 0.25 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 23 | 3 | 3 | 1.667 | 100 | 0.6 | 0 | 0.333 | 33.333 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 24 | 7 | 7 | 7.5 | 100 | 1.4 | 0 | 0.429 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 25 | 6 | 6 | 9.167 | 100 | 1.2 | 0 | 0.667 | 16.667 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 26 | 1 | 1 | 0.833 | 100 | 0.2 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 27 | 6 | 6 | 6.667 | 100 | 1.2 | 0 | 1.167 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 28 | 4 | 4 | 6.667 | 100 | 0.8 | 0 | 2.25 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 29 | 3 | 3 | 5 | 100 | 0.6 | 0 | 1.667 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 30 | 6 | 6 | 4.167 | 100 | 1.2 | 0 | 0.667 | 16.667 | 0 | 0 | 0 | 0 |
| 2013 | 1 | 31 | 5 | 5 | 5.833 | 100 | 1 | 0 | 4.6 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 1 | 4 | 4 | 4.167 | 100 | 0.8 | 0 | 3 | 25 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 2 | 3 | 3 | 4.167 | 100 | 0.6 | 0 | 0.667 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 3 | 4 | 4 | 5 | 100 | 0.8 | 0 | 1.5 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 4 | 6 | 6 | 5 | 100 | 1.2 | 0 | 1.5 | 16.667 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 5 | 7 | 7 | 5.833 | 100 | 1.4 | 0 | 0.714 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 6 | 7 | 7 | 9.167 | 100 | 1.4 | 0 | 3.571 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 7 | 9 | 9 | 30.833 | 100 | 1.8 | 0 | 1.111 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 8 | 6 | 6 | 4.167 | 100 | 1.2 | 0 | 0.5 | 16.667 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 9 | 5 | 5 | 5.833 | 100 | 1 | 0 | 1.2 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 10 | 2 | 2 | 0.833 | 100 | 0.4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

| Año | Mes | Día | NIQ | NPO | ASO | GCOP | RQ | PIQS | PDP | PIPQ | PCI | PCA | PCR | TM |
|------|-----|-----|-----|-----|--------|------|-----|------|-------|--------|-----|-----|-----|----|
| 2013 | 2 | 11 | 4 | 4 | 2.5 | 100 | 0.8 | 0 | 2.75 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 12 | 3 | 3 | 2.5 | 100 | 0.6 | 0 | 1.333 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 13 | 5 | 5 | 6.667 | 100 | 1 | 0 | 0.6 | 20 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 14 | 7 | 7 | 8.333 | 100 | 1.4 | 0 | 0.429 | 14.286 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 15 | 3 | 3 | 4.167 | 100 | 0.6 | 0 | 0.333 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 16 | 1 | 1 | 0.833 | 100 | 0.2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 17 | 4 | 4 | 5 | 100 | 0.8 | 0 | 0.5 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 18 | 3 | 3 | 1.667 | 100 | 0.6 | 0 | 1.333 | 33.333 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 19 | 5 | 5 | 5.833 | 100 | 1 | 0 | 0.2 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 20 | 7 | 7 | 9.167 | 100 | 1.4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 21 | 10 | 10 | 10 | 100 | 2 | 0 | 0.6 | 10 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 22 | 8 | 8 | 10 | 100 | 1.6 | 0 | 1.375 | 25 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 24 | 4 | 4 | 3.333 | 100 | 0.8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 25 | 8 | 8 | 16.667 | 100 | 1.6 | 0 | 0.375 | 37.5 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 26 | 9 | 8 | 6.667 | 100 | 1.8 | 0 | 1.125 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 27 | 5 | 5 | 16.667 | 100 | 1 | 0 | 5.8 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 2 | 28 | 8 | 8 | 3.333 | 100 | 1.6 | 0 | 0.125 | 75 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 1 | 7 | 7 | 5.833 | 100 | 1.4 | 0 | 1.857 | 28.571 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 2 | 2 | 2 | 0.833 | 100 | 0.4 | 0 | 0.5 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 3 | 1 | 2 | 5 | 50 | 0.2 | 50 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 4 | 8 | 8 | 20 | 100 | 1.6 | 0 | 1.625 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 5 | 6 | 6 | 7.5 | 100 | 1.2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 6 | 7 | 7 | 28.333 | 100 | 1.4 | 0 | 0 | 14.286 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 7 | 7 | 7 | 8.333 | 100 | 1.4 | 0 | 0.286 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 8 | 5 | 5 | 6.667 | 100 | 1 | 0 | 1.6 | 0 | 0 | 0 | 0 | 0 |

| Año | Mes | Día | NIQ | NPO | ASO | GCOP | RQ | PIQS | PDP | PIPQ | PCI | PCA | PCR | TM |
|------|-----|-----|-----|-----|--------|------|-----|------|-------|--------|-----|-----|-----|----|
| 2013 | 3 | 8 | 5 | 5 | 6.667 | 100 | 1 | 0 | 1.6 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 9 | 4 | 4 | 21.667 | 100 | 0.8 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 10 | 5 | 5 | 3.333 | 100 | 1 | 0 | 0.2 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 11 | 9 | 9 | 7.5 | 100 | 1.8 | 0 | 0.444 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 12 | 7 | 7 | 19.167 | 100 | 1.4 | 0 | 0.857 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 13 | 8 | 8 | 10.833 | 100 | 1.6 | 0 | 1.875 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 14 | 8 | 8 | 8.333 | 100 | 1.6 | 0 | 1.25 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 15 | 3 | 3 | 8.333 | 100 | 0.6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 16 | 3 | 3 | 3.333 | 100 | 0.6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 18 | 7 | 7 | 10.833 | 100 | 1.4 | 0 | 5.286 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 19 | 4 | 4 | 6.667 | 100 | 0.8 | 0 | 0.75 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 20 | 5 | 5 | 7.5 | 100 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 21 | 11 | 11 | 15 | 100 | 2.2 | 0 | 1.545 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 22 | 4 | 4 | 1.667 | 100 | 0.8 | 0 | 0.75 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 23 | 1 | 1 | 1.667 | 100 | 0.2 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 25 | 9 | 9 | 11.667 | 100 | 1.8 | 0 | 0.222 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 26 | 6 | 6 | 5.833 | 100 | 1.2 | 0 | 0.667 | 16.667 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 27 | 2 | 2 | 15.833 | 100 | 0.4 | 0 | 6 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 28 | 3 | 3 | 2.5 | 100 | 0.6 | 0 | 2.667 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 29 | 6 | 6 | 10 | 100 | 1.2 | 0 | 2.5 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 3 | 31 | 5 | 5 | 5 | 100 | 1 | 0 | 1.2 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 1 | 8 | 8 | 19.167 | 100 | 1.6 | 0 | 4.625 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 2 | 8 | 8 | 8.333 | 100 | 1.6 | 0 | 1.625 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 3 | 4 | 4 | 20 | 100 | 0.8 | 0 | 0.5 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 4 | 6 | 6 | 6.667 | 100 | 1.2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

| Año | Mes | Día | NIQ | NPO | ASO | GCOP | RQ | PIQS | PDP | PIPQ | PCI | PCA | PCR | TM |
|------|-----|-----|-----|-----|--------|------|-----|------|-------|------|--------|-----|-----|--------|
| 2013 | 4 | 5 | 10 | 10 | 9.167 | 100 | 2 | 0 | 0.6 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 6 | 7 | 7 | 5 | 100 | 1.4 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 7 | 2 | 2 | 3.333 | 100 | 0.4 | 0 | 0.5 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 8 | 3 | 3 | 3.333 | 100 | 0.6 | 0 | 3.333 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 9 | 7 | 7 | 19.167 | 100 | 1.4 | 0 | 1.286 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 10 | 10 | 10 | 4.167 | 100 | 2 | 0 | 0.2 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 11 | 8 | 8 | 8.333 | 100 | 1.6 | 0 | 0.375 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 12 | 4 | 4 | 9.167 | 100 | 0.8 | 0 | 0.5 | 25 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 13 | 4 | 4 | 5 | 100 | 0.8 | 0 | 1.5 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 14 | 4 | 4 | 7.5 | 100 | 0.8 | 0 | 0.25 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 15 | 4 | 4 | 3.333 | 100 | 0.8 | 0 | 0.5 | 25 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 16 | 6 | 6 | 18.333 | 100 | 1.2 | 0 | 1.333 | 0 | 16.667 | 0 | 0 | 16.667 |
| 2013 | 4 | 17 | 2 | 2 | 3.333 | 100 | 0.4 | 0 | 6.5 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 18 | 5 | 5 | 15 | 100 | 1 | 0 | 8.8 | 0 | 0 | 0 | 0 | 20 |
| 2013 | 4 | 19 | 7 | 7 | 9.167 | 100 | 1.4 | 0 | 0.857 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 20 | 6 | 6 | 7.5 | 100 | 1.2 | 0 | 0.667 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 21 | 1 | 1 | 0 | 100 | 0.2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 22 | 1 | 1 | 1.667 | 100 | 0.2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 23 | 6 | 6 | 6.667 | 100 | 1.2 | 0 | 1.333 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 24 | 3 | 3 | 3.333 | 100 | 0.6 | 0 | 0.333 | 0 | 0 | 0 | 0 | 0 |



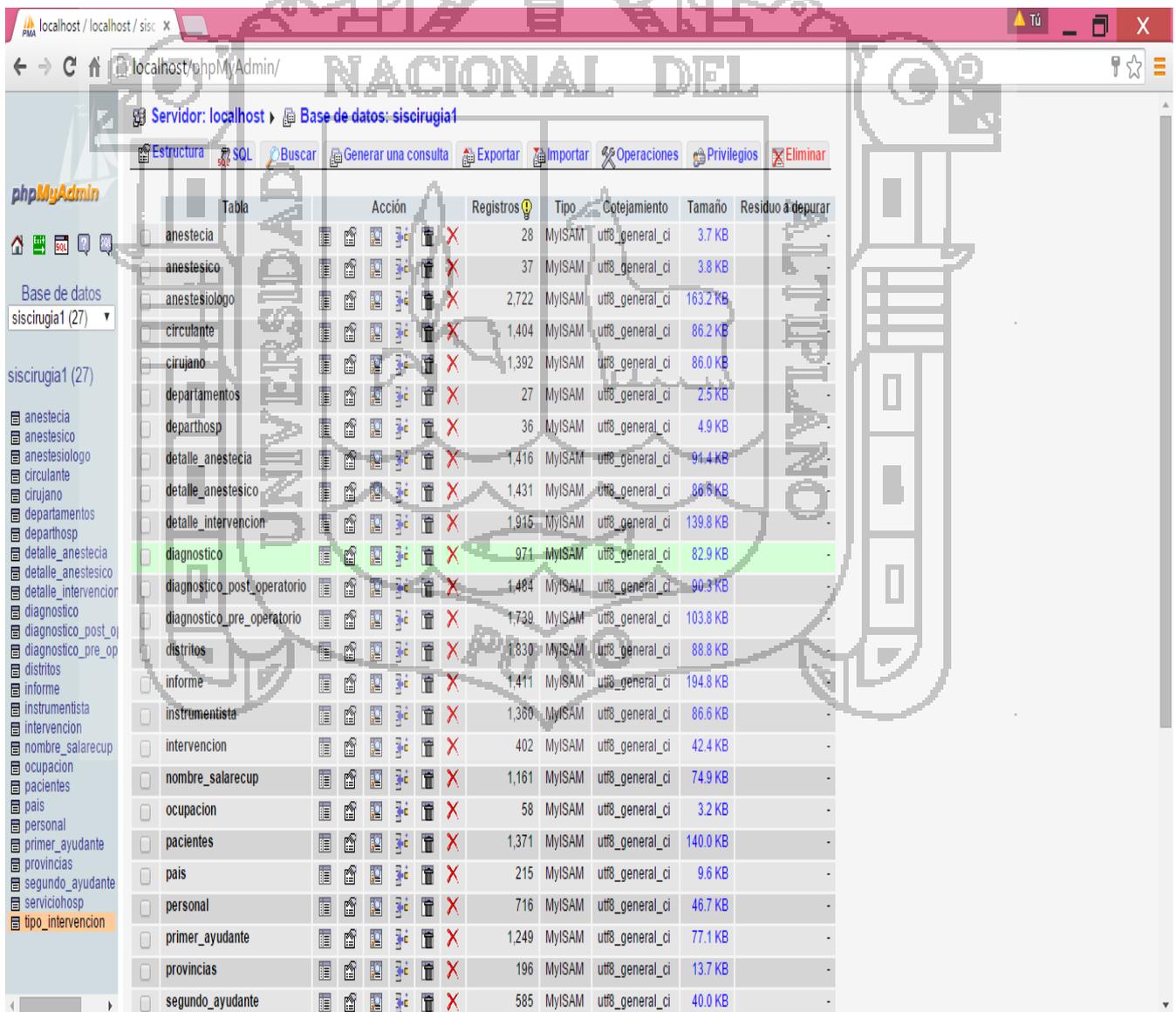
| Año | Mes | Día | NIQ | NPO | ASO | GCOP | RQ | PIQS | PDP | PIPQ | PCI | PCA | PCR | TM |
|------|-----|-----|-----|-----|--------|------|-----|------|-------|------|-----|-----|-----|----|
| 2013 | 4 | 25 | 3 | 3 | 3.333 | 100 | 0.6 | 0 | 0.667 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 26 | 7 | 7 | 40.833 | 100 | 1.4 | 0 | 0.286 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 27 | 6 | 6 | 7.5 | 100 | 1.2 | 0 | 1.333 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 28 | 3 | 3 | 5 | 100 | 0.6 | 0 | 2.667 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 29 | 8 | 8 | 9.167 | 100 | 1.6 | 0 | 0.125 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 4 | 30 | 6 | 6 | 6.667 | 100 | 1.2 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 5 | 1 | 2 | 2 | 2.5 | 100 | 0.4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 5 | 2 | 6 | 6 | 7.5 | 100 | 1.2 | 0 | 0.333 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 5 | 3 | 3 | 3 | 1.667 | 100 | 0.6 | 0 | 0.667 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 5 | 4 | 4 | 4 | 3.333 | 100 | 0.8 | 0 | 0.25 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 5 | 5 | 1 | 1 | 0 | 100 | 0.2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 5 | 6 | 4 | 5 | 4.167 | 80 | 0.8 | 20 | 4.2 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 5 | 7 | 0 | 8 | 5.833 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 5 | 8 | 0 | 2 | 3.333 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 5 | 9 | 0 | 7 | 17.5 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 5 | 10 | 0 | 5 | 7.5 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 5 | 11 | 0 | 1 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 5 | 12 | 0 | 1 | 0.833 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 5 | 13 | 0 | 6 | 8.333 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2013 | 5 | 14 | 0 | 1 | 0.833 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |

ANEXO 6: PROCESO ETL

Para realizar el proceso de ETL lo primero que se realizó fue instalar el motor de base de datos utilizado que es MYSQL y el Software de Minería de datos Weka.

1. Extracción de los datos

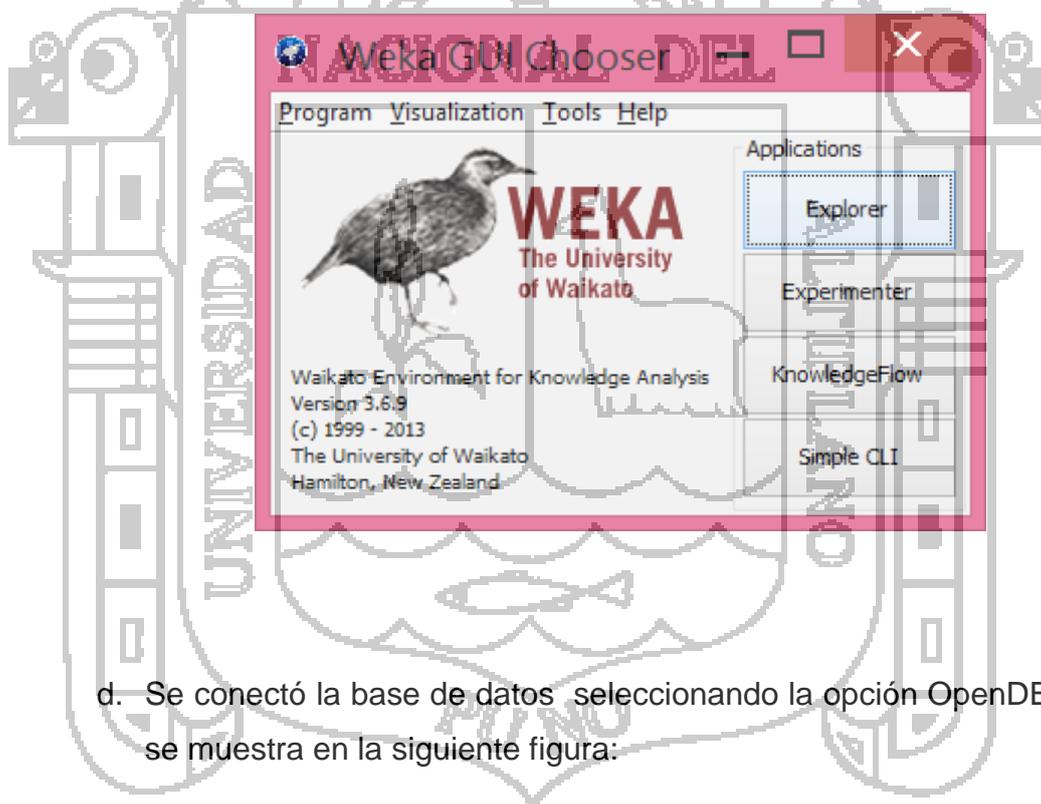
Se extrajo los datos de la base de datos siscirugia1 almacenados en el motor de base de datos de MYSQL:



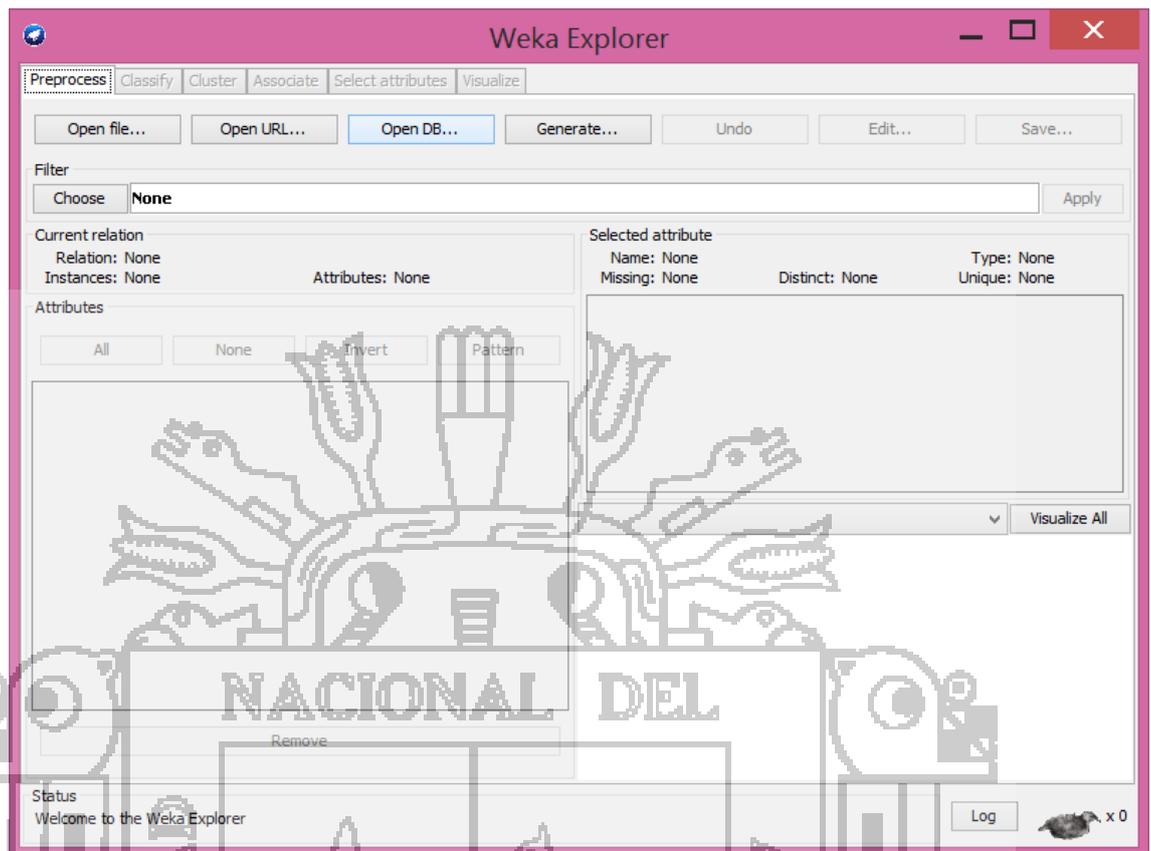
| Tabla | Acción | Registros | Tipo | Cotejamiento | Tamaño | Residuo a depurar |
|-----------------------------|--------|-----------|--------|-----------------|----------|-------------------|
| anestecia | | 28 | MyISAM | utf8_general_ci | 3.7 KB | |
| anestesico | | 37 | MyISAM | utf8_general_ci | 3.8 KB | |
| anestesiologo | | 2,722 | MyISAM | utf8_general_ci | 163.2 KB | |
| circulante | | 1,404 | MyISAM | utf8_general_ci | 86.2 KB | |
| cirujano | | 1,392 | MyISAM | utf8_general_ci | 86.0 KB | |
| departamentos | | 27 | MyISAM | utf8_general_ci | 2.5 KB | |
| departhosp | | 36 | MyISAM | utf8_general_ci | 4.9 KB | |
| detalle_anestecia | | 1,416 | MyISAM | utf8_general_ci | 91.4 KB | |
| detalle_anestesico | | 1,431 | MyISAM | utf8_general_ci | 86.5 KB | |
| detalle_intervencion | | 1,915 | MyISAM | utf8_general_ci | 139.8 KB | |
| diagnostico | | 971 | MyISAM | utf8_general_ci | 82.9 KB | |
| diagnostico_post_operatorio | | 1,484 | MyISAM | utf8_general_ci | 90.3 KB | |
| diagnostico_pre_operatorio | | 1,739 | MyISAM | utf8_general_ci | 103.8 KB | |
| distritos | | 1,830 | MyISAM | utf8_general_ci | 88.8 KB | |
| informe | | 1,411 | MyISAM | utf8_general_ci | 194.8 KB | |
| instrumentista | | 1,360 | MyISAM | utf8_general_ci | 86.6 KB | |
| intervencion | | 402 | MyISAM | utf8_general_ci | 42.4 KB | |
| nombre_salarecup | | 1,161 | MyISAM | utf8_general_ci | 74.9 KB | |
| ocupacion | | 58 | MyISAM | utf8_general_ci | 3.2 KB | |
| pacientes | | 1,371 | MyISAM | utf8_general_ci | 140.0 KB | |
| pais | | 215 | MyISAM | utf8_general_ci | 9.6 KB | |
| personal | | 716 | MyISAM | utf8_general_ci | 46.7 KB | |
| primer ayudante | | 1,249 | MyISAM | utf8_general_ci | 77.1 KB | |
| provincias | | 196 | MyISAM | utf8_general_ci | 13.7 KB | |
| segundo ayudante | | 585 | MyISAM | utf8_general_ci | 40.0 KB | |

Para extraer los datos de MYSQL en Weka lo que se realizó fue:

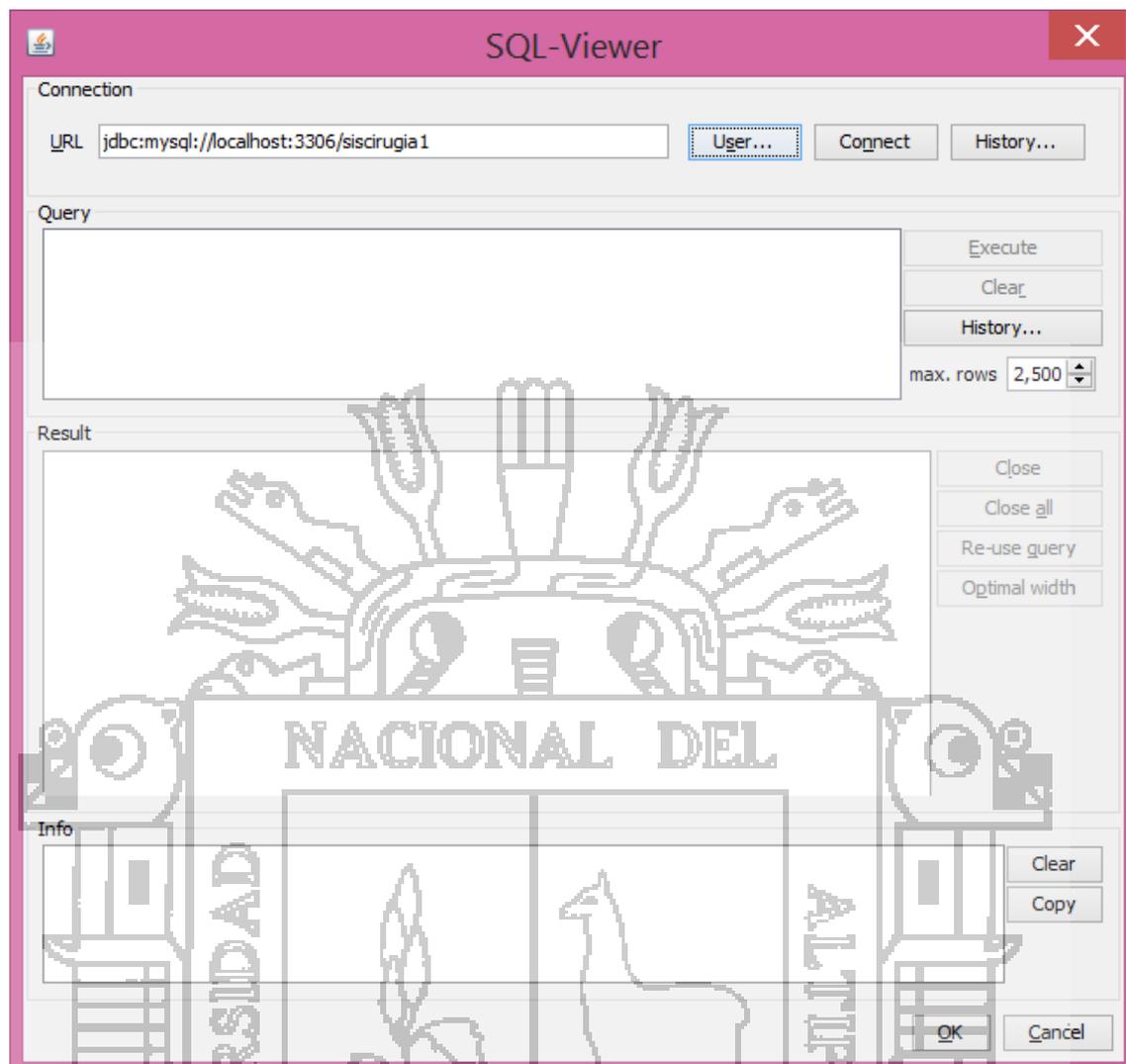
- a. Descargar mysql-conector-java-5.1.24.zip que es el driver de esta carpeta se copió el archivo mysql-conector-java-5.1.24-bin.jar en la ruta donde se instaló Weka que es C:\Program Files\Weka-3-6.
- b. En la ruta C:\Program Files\Weka-3-6 se encuentra un archivo de texto RunWeka.ini se ejecutó como administrador y editó la última línea ubicando la ruta siguiente cp=%CLASSPATH%;mysql-conector-java-5.1.24-bin.jar.
- c. Se inició el programa WEKA seleccionando la opción Explorer.



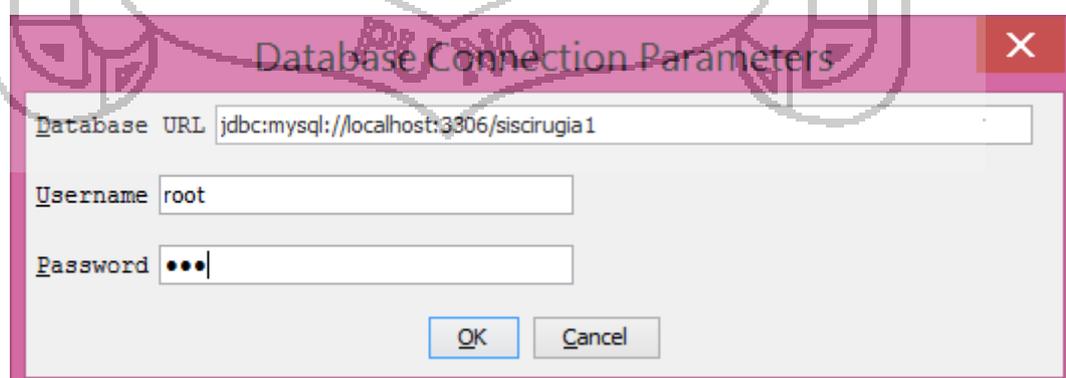
- d. Se conectó la base de datos seleccionando la opción OpenDB como se muestra en la siguiente figura:



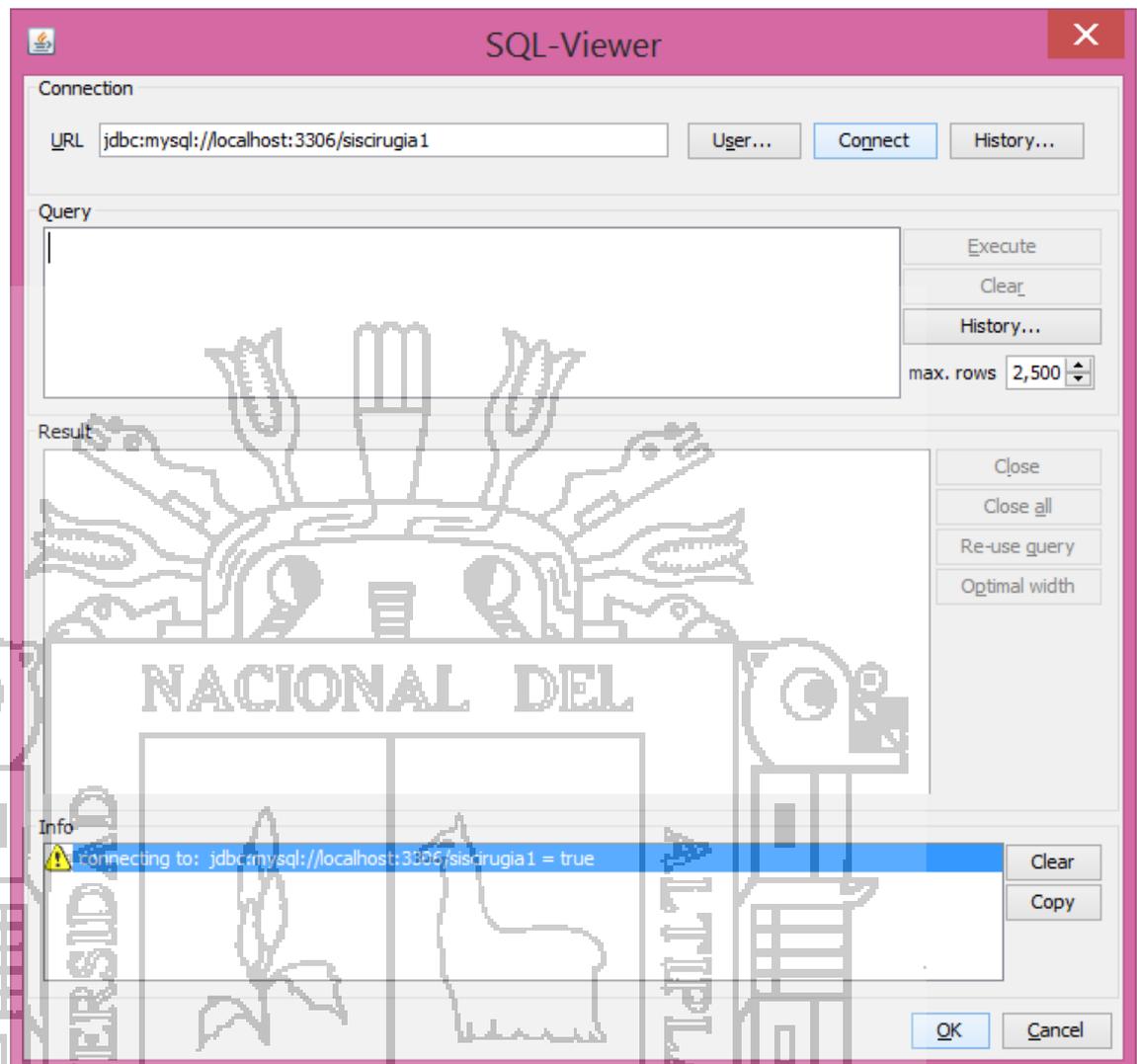
e. Seguidamente se editó la URL en el que se colocó la ruta de la base de datos como se muestra continuación en la siguiente figura:



- f. Luego se configuró el usuario y contraseña que se le colocó a la base de datos de MSQL.



- g. Para verificar la conexión se seleccionó la opción connect



2. Transformación de los datos

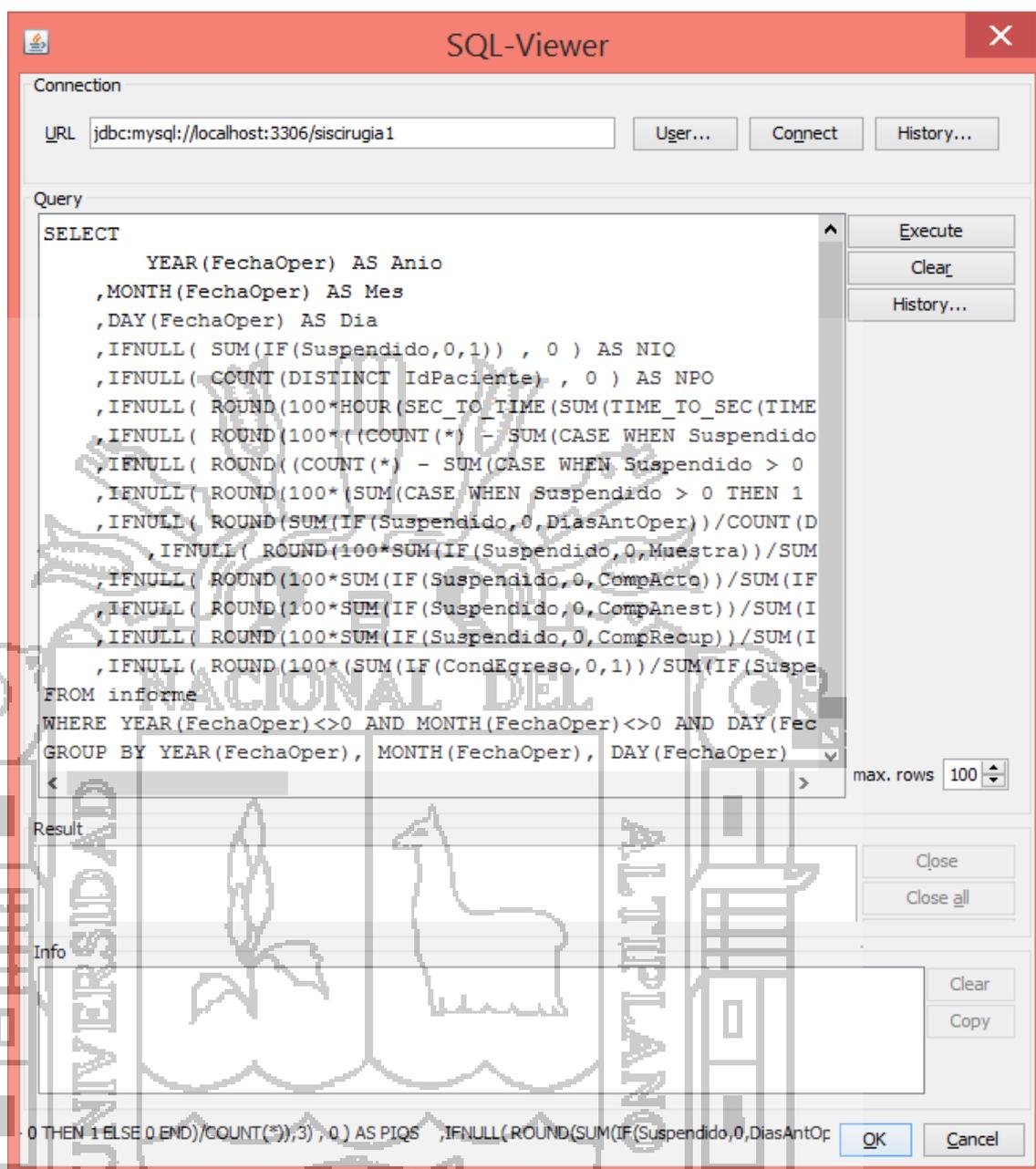
A continuación se muestra la consulta para obtener los datos con los que se trabajó:

```

SELECT
    YEAR(FechaOper) AS Anio
    ,MONTH(FechaOper) AS Mes
    ,DAY(FechaOper) AS Dia
    ,IFNULL( SUM(IF(Suspendido,0,1)) , 0 ) AS NIQ
    ,IFNULL( COUNT(DISTINCT IdPaciente) , 0 ) AS NPO
    ,IFNULL(
ROUND(100*HOUR(SEC_TO_TIME(SUM(TIME_TO_SEC(TIMEDIFF(IF(TIMEDIFF(HoraFinOper,HoraIngr
esoSala)<0,ADDTIME(HoraFinOper,'24:00:00'),HoraFinOper),HoraIngresoSala)))))/(5*24),3) , 0 ) AS
ASO
    ,IFNULL( ROUND(100*((COUNT(*) - SUM(CASE WHEN Suspendido > 0 THEN 1 ELSE 0
END))/COUNT(*)),3) , 0 ) AS GCOP
    ,IFNULL( ROUND((COUNT(*) - SUM(CASE WHEN Suspendido > 0 THEN 1 ELSE 0 END))/5,3) , 0 ) AS
RQ
    ,IFNULL( ROUND(100*(SUM(CASE WHEN Suspendido > 0 THEN 1 ELSE 0 END)/COUNT(*)),3) , 0 ) AS
PIQS
    ,IFNULL( ROUND(SUM(IF(Suspendido,0,DiasAntOper))/COUNT(DISTINCT IdPaciente),3) , 0 ) AS PDP
    ,IFNULL( ROUND(100*SUM(IF(Suspendido,0,Muestra))/SUM(IF(Suspendido,0,1)),3) , 0 ) AS
PIPQ
    ,IFNULL( ROUND(100*SUM(IF(Suspendido,0,CompActo))/SUM(IF(Suspendido,0,1)),3) , 0 ) AS PCI
    ,IFNULL( ROUND(100*SUM(IF(Suspendido,0,CompAnest))/SUM(IF(Suspendido,0,1)),3) , 0 ) AS PCA
    ,IFNULL( ROUND(100*SUM(IF(Suspendido,0,CompRecup))/SUM(IF(Suspendido,0,1)),3) , 0 ) AS PCR
    ,IFNULL( ROUND(100*(SUM(IF(CondEgreso,0,1))/SUM(IF(Suspendido,0,1))),3) , 0 ) AS TM
FROM informe
WHERE YEAR(FechaOper)<>0 AND MONTH(FechaOper)<>0 AND DAY(FechaOper)<>0
GROUP BY YEAR(FechaOper), MONTH(FechaOper), DAY(FechaOper)

```

a. En weka en la sección query se realizó la consulta a utilizar seleccionando los datos para los 12 indicadores Hospitalarios utilizando las formulas de cada uno obteniendo los resultados deseados como se muestra en la siguiente figura:



b. Seguidamente se ejecutó en Weka la consulta anterior obteniendo los datos transformados que se utilizaron como se muestra a continuación en la figura:

SQL-Viewer

Connection
 URL: User... Connect History...

Query

```

SELECT
    YEAR(FechaOper) AS Anio
    ,MONTH(FechaOper) AS Mes
    ,DAY(FechaOper) AS Dia
    ,IFNULL( SUM(IF(Suspendido,0,1)) , 0 ) AS NIQ
    ,IFNULL( COUNT(DISTINCT IdPaciente) , 0 ) AS NPO
    ,IFNULL( ROUND(100*HOUR(SEC_TO_TIME(SUM(TIME_TO_SEC(TIM
    ,IFNULL( ROUND(100*((COUNT(*) - SUM(CASE WHEN Suspendido
    ,IFNULL( ROUND((COUNT(*) - SUM(CASE WHEN Suspendido > 0
    ,IFNULL( ROUND(100*(SUM(CASE WHEN Suspendido > 0 THEN 1
    ,IFNULL( ROUND(SUM(IF(Suspendido,0,DiasAntOper))/COUNT(
    ,IFNULL( ROUND(100*SUM(IF(Suspendido,0,Muestra))/SU
    ,IFNULL( ROUND(100*SUM(IF(Suspendido,0,CompActo))/SUM(I
    ,IFNULL( ROUND(100*SUM(IF(Suspendido,0,CompAnest))/SUM(
    ,IFNULL( ROUND(100*SUM(IF(Suspendido,0,CompRecup))/SUM(
    ,IFNULL( ROUND(100*(SUM(IF(CondEgreso,0,1))/SUM(IF(Susp
    FROM informe
    WHERE YEAR(FechaOper)<>0 AND MONTH(FechaOper)<>0 AND DAY(Fe
    GROUP BY YEAR(FechaOper), MONTH(FechaOper), DAY(FechaOper)
    
```

 Execute Clear History... max. rows 2,500

Result

| Row | Anio | Mes | Dia | NIQ | NPO | ASO | GCOP | RQ | PIQS | PDP | PIPQ | PCI | PCA | Close |
|--------|------|-----|-----|-----|-----|-----|------|----|------|-----|------|-----|-----|-----------|
| Query1 | | | | | | | | | | | | | | Close all |

Info
 connecting to: jdbc:mysql://localhost:3306/siscirugia1 = true
 Query: SELECT YEAR(FechaOper) AS Anio ,MONTH(FechaOper) AS Mes ,DAY(FechaOper) AS Dia ,IFNULL
 316 rows selected.
 Clear Copy

0 THEN 1 ELSE 0 END)/COUNT(*)),3) , 0.) AS PIQS ,IFNULL(ROUND(SUM(IF(Suspendido,0,DiasAntOp

OK Cancel

3. Carga de los datos

Finalmente para la carga de los datos se selecciona ok y se cargan todos los datos transformados en el Software Weka.

